

الإحصاء والاحتمالات في التطبيقات الهندسية

الإحصاء والإحتمالات فى التطبيقات الهندسية

Statistics and Probability in Engineering Applications

د. أمجد إبراهيم شحادة د. علي إبراهيم سعد

م. محمد رياض علي

دار الفجر للنشر والتوزيع

2005

الإحصاء والاحتمالات فى التطبيقات الهندسية

Statistics and Probability in Engineering Applications

د. أمجد إبراهيم شحادة د. على إبراهيم سعد

د. محمد رياض على

رقم الإيداع 7705 الترقيم الدولي. I.S.B.N. 977-358-093-8	حقوق النشر الطبعة الأولى 2005 م جميع الحقوق محفوظة للناسر
--	---

دار الفجر للنشر و التوزيع

4 شارع هاشم الأشقر - الزهة الجديدة - القاهرة

ت : 6242520 - 6246252 (00202)

ف : 6246265 (00202)

لا يجوز نشر أي جزء من الكتاب أو اختزان مادته بطريقة الاسترجاع أو نقله على أي نحو أو بأي طريقة سواء كانت إلكترونية أو ميكانيكية أو بخلاف ذلك إلا بموافقة الناسر على هذا كتابة و مقدما .

بسم الله الرحمن الرحيم

﴿ وأحاط بما لديهم وأحصى كل شيء عدداً ﴾

الإهداء

إلى دارسي الهندسة ... وإلى كل الذين يترقبون الأمل بنجاحنا وإلى
كل الذين لمسنا منهم العون والتشجيع في أنجاز هذا العمل أجلاً
واحتراماً آملي أن يكون هذا العمل إضافة جديدة إلى المكتبة العربية .

المؤلفون

محتويات الكتاب

7 : مقدمة الكتاب
13 : الباب الأول : مدخل في علم الإحصاء
67 : الباب الثاني : مقاييس النزعة المركزية
152 : الباب الثالث : مقاييس التشتت
196 : الباب الرابع : الارتباط والانحدار
256 : الباب الخامس : مبادئ نظرية الاحتمالات
351 : الباب السادس : المتغيرات العشوائية والتوزيعات الاحتمالية
408 : جداول التوزيعات الاحتمالية المختلفة
415 : المراجع

يعد علم الإحصاء (Statistics) احد أقدم العلوم الطبيعية ، حيث يذهب المؤرخون إلى أن الإنسان كان قد لجأ إلى جمع البيانات الإحصائية وأستخدم الإحصاءات منذ أقدم العصور ، حيث كانت تجمع بيانات عن عدد السكان والنشاطات الزراعية والاقتصادية ، فقد حرص الصينيون القدماء والفراعنة المصريون وغيرهم من الحضارات القديمة على الاحتفاظ بسجلات إحصاء عدد السكان وممتلكاتهم ، وذلك بهدف تقدير حجم المحاصيل الزراعية لتسعيروها وفرض وتقدير الضرائب عليها .

وقد ورد ذكر علم الإحصاء في القرآن الكريم بنفس الغرض الذي يستخدم به هذا العلم الآن وهو الحصر والعد مثل قوله تعالى: " وأن تعدوا نعمة الله لا تحصوها " ، وقوله تعالى : " وأحاط بما لديهم وأحصى كل شيء عدداً " .

أن تاريخ العلوم يؤكد أن الحضارة الحديثة تدين بازدهارها أساساً للحضارة العربية الإسلامية بما نقلت عنها من أصول العلم وتفرعاته . فقد كان للمسلمين الدور الواضح والبارز في استخدام الكثير من الإحصاءات الخاصة بتنظيم الدولة الإسلامية عندما كانت في أوجه ازدهارها ، وعبر التاريخ وحتى الوقت الحاضر زادت أهمية الإحصاءات وأصبحت مألوفة لدينا وتمثل جانباً هاماً من المعلومات التي نطالعها كل يوم .

من ذلك على سبيل المثال إتجازات دولة ما في مجال الإسكان والأمن الغذائي ، والتقديرات الخاصة بالتنبؤات الجوية ، وإحصاءات النقاط التي تحرزها أندية كرة القدم ، والتغيرات التي تطرأ على أسعار العملات وغيرها .

ولقد أصبح التخطيط في الوقت الحاضر أسلوباً علمياً وطريقة تعتمد عليها الدول في إداراتها الصناعية والاقتصادية والاجتماعية المختلفة . وأدركت الكثير من الدول الحديثة وإداراتها الحكومية والخاصة بعد ممارستها لهذا الأسلوب في تنظيم وإدارة أعمالها أنها بحاجة إلى الإحصاءات الدقيقة عن المتغيرات المختلفة التي تركز عليها برامج العمل والإنتاج ، حتى يمكن أعدادها أعداداً صحيحاً ، ومتابعة تنفيذها للكشف عن جوانب الضعف التي قد تعترض مجرى العمليات المتشابهة التي تتضمنها هذه البرامج ، وتقييم النتائج النهائية بعد انتهاء مراحل التنفيذ ، وذلك من أجل استخلاص النتائج والتوجيهات والتي تعود بالفائدة عند إعداد الخطط المستقبلية المختلفة .

وقد أدت هذه الحاجة الضرورية إلى الإحصاءات الدقيقة عن مختلف النواحي والمجالات إلى اهتمام المسؤولين في معظم الدول وفي جميع القطاعات العامة والخاصة بتدريس طرق الإحصاء وأساليب التحليل الإحصائي ، وربط تلك الأساليب بنظرية الاحتمالات (Thoery of Probability) ، لما لذلك من أهمية بالغة في الوقت الراهن في أغلب اتجاهات العلوم الهندسية والتطبيقية المختلفة . لذلك تزايد الطلب في الوقت الحاضر على إنشاء معاهد للتدريس والتدريب الإحصائي في جميع دول العالم وذلك لتأهيل الكوادر على القيام بالإعمال الإحصائية المختلفة .

ومما لا شك فيه أن الدول العربية جميعاً في هذه المرحلة من تطورها هي في أشد الحاجة إلى الانفتاح على هذا الأسلوب في تنظيم وإدارة الأعمال وبالأذات في حقل الأبحاث الهندسية التطبيقية المرتبطة في جميع النشاطات الصناعية والاقتصادية والاجتماعية ، حتى تستطيع أن تواجه التطور الحضاري الحديث ، وأن تأخذ مكانها اللائق بها بين الدول المتقدمة في مجالات الأبحاث العلمية ذات المستوى الرفيع .

في الوقت الحالي ارتبط علم الإحصاء بنظرية الاحتمالات ، التي ترجع بداية ظهورها إلى منتصف القرن السابع عشر كنتيجة لدراسة بعض ألعاب الحظ المختلفة . ومنذ ذلك الوقت أشترك في أبحاث نظرية الاحتمالات الكثير من العلماء أمثال هيوجنس وباسكال وفيرما وجاكوب بونولي وغيرهم . لكن متطلبات العلوم الطبيعية والتطبيقات الاجتماعية مثل مشاكل الإحصاء وخاصة الأحصاء السكاني ، ونظرية أخطاء المشاهدات أدت إلى ضرورة الاستمرار في تطوير نظرية الاحتمالات .

وقد لعب كل من العلماء أمثال لابلاس وجولس وبواسون دوراً هاماً في تطوير الطرق التحليلية لنظرية الاحتمالات في منتصف القرن الثامن عشر ، وقد ساهم في تطوير نظرية الاحتمالات منذ منتصف القرن التاسع عشر وحتى العشرينات تقريباً من القرن العشرين وبدرجة كبيرة العلماء الروس أمثال (تشيبيسوف ، وماركوف ، وليبابونوف) الذين شجعوا بشكل واسع على انتشار البحوث المتعلقة بربط واستخدام نظرية الاحتمالات في الأحصاء وخاصة في أمور التأمين والإحصاء السكاني ، حيث تنحصر الأهمية البالغة لأعمال هؤلاء العلماء في أنهم أدخلوا مفهوم المقادير العشوائية كمادة للدراسة المستمرة واستخدموه بشكل واسع مما مهد الطريق إلى توأمة هذه النظرية بعلم الأحصاء وارتباطها الوثيق به .

إلا أن هذه النظرية لم توضع لها مسلمات إلا في الثلاثينات من القرن العشرين . وأصبحت تعرف على أنها العلم الذي يدرس الظواهر العشوائية ، وقد تطورت نظرية الاحتمالات تلبية لمتطلبات الحياة العملية ، مثلها مثل أجزاء العلوم الرياضية الأخرى ، حيث أن العلاقة المبنية بين نظرية الاحتمالات ومتطلبات العلوم الطبيعية ، توضح بأفضل ما يمكن تلك الأسباب التي جعلت نظرية الاحتمالات في العقود الأخيرة من أسرع

فروع الرياضيات تطوراً . فالنتائج النظرية الجديدة تعمل على فتح آفاقاً جديدة لاستعمال طرق نظرية الاحتمالات في العلوم الطبيعية ، وتقوم الدراسة الشاملة لظواهر الطبيعة بدفع نظرية الاحتمالات إلى الكشف عن قوانين جديدة ولدت بالصدفة .

وقد كبرت أهمية ارتباط علم الاحصاء بنظرية الاحتمالات في السنوات الأخيرة بفعل التطور التقني والصناعي السريع . وكنتيجة لذلك فقد تعاظم الاهتمام بنتائج نظرية الاحتمالات لا لفرز البضاعة التي تم إنتاجها فحسب بل والاهم من ذلك لتنظيم عملية الإنتاج ذاتها ، والرقابة الإحصائية المتعلقة بمشاكل التحقق من نوعية المنتجات ، وبالتالي ظهرت نظرية الطرق الإحصائية لرقابة القبول العميقة بمحتواها ، والهامة بتطبيقاتها العملية والمبنية على الاستخدام الواسع لنظرية الاحتمالات .

أن لوضع طرق إحصائية للتحكم بنوعية المنتجات خلال عملية الإنتاج أهمية كبيرة جداً في حلقة هذه الأفكار . حيث أصبحت نظرية الوثوق (Theory of Reliability) ، التي تستخدم بصورة واسعة طرق نظرية الاحتمالات تلعب دوراً هاماً في جميع العلوم الهندسية . مما أدى إلى تداخل أفكار وقوانين نظرية الاحتمالات مع علم الإحصاء في أغلب الاتجاهات مثل الطبيعة والكيمياء والطب وعلم النفس وإدارة الأعمال والاقتصاد والهندسة بمختلف فروعها .

أن دراسة علم الاحصاء ونظرية الاحتمالات يعتبر أمر هاماً وذو فوائد كثيرة بالنسبة لدارسي الهندسة بمختلف فروعها ، وذلك نظراً لارتباطه الوثيق بالعلوم التطبيقية الهندسية المختلفة ، وخاصة في مجال الأبحاث الهندسية والدراسات العليا للمهندسين ، وبذات بعد أن تفتحت أمامهم في الوقت الحالي مجالات عمل كثيرة في الشركات العامة والخاصة ومراكز

البحوث وغيرها . بل أن المعرفة بعلم الاحصاء والاحتمالات تفيد الإنسان على المستوى الشخصي فتكسبه مهارة التخطيط لحياته الاقتصادية العامة .

أن القليل من الطلبة دارسي مادة الاحصاء والاحتمالات ، ينظرون إلى هذه المادة باعتبارها موضوعاً يثير الاهتمام والدراسة الموسعة مقارنة مع فروع العلوم الرياضية الأخرى ، أما معظم الطلاب فيرون أن موضوع الاحصاء والاحتمالات تكتنفه الصعوبات والتعقيدات ويستعصى على الفهم والاستيعاب ، ولهذه الأسباب وغيرها أقدمنا على تقديم هذا الكتاب بأسلوب سهل ومبسط لدارسي الهندسة بمختلف فروعها .

وقد تركز الاهتمام في هذا الكتاب على أمداد الطلبة بالمبادئ الإحصائية الضرورية وتبويب البيانات ، وقياس علاقة الارتباط بين المتغيرات المختلفة ، وربطها بمبادئ نظرية الاحتمالات ، وقد روعي في عرض مادة الكتاب أهمية ربطها بالتطبيقات الهندسية المختلفة التي يهتم بها دارسو الهندسة بمختلف فروعها .

أن الهدف الرئيسي لهذا الكتاب يكمن في تقديم المبادئ والمقاييس الأساسية لعلم الإحصاء ونظرية الاحتمالات ببساطة ووضوح ، وتوضيح كيفية التعامل مع البيانات من حيث تبويبها وعرضها بيانياً وأجراء العمليات المختلفة عليها ، بالإضافة إلى دراسة أساليب الإحصاء الاستنتاجي التي تساعد على فحص النتائج ومقارنتها بطريقة علمية منظمة . وقياس علاقة الارتباط بين المتغيرات المختلفة ، وربط كل ما ذكر بالتطبيقات الهندسية المختلفة من خلال الاستعانة بالعديد من الأمثلة التوضيحية المحولة والتمارين المنتقاة بعناية . وقد حاولنا قدر الامكان أن تكون أبواب الكتاب متناسقة حتى تكون عملية الدراسة مريحة ومثمرة .

أن المشكلة الحقيقية المطروحة أمام العلم هي تطوير وتحسين الإنتاج وبشكل سريع ، واعتماد التخطيط كمنهج وأسلوب علمي في الإدارات الصناعية والاقتصادية والاجتماعية ، والحاجة الماسة إلى الإحصاءات الدقيقة عن المتغيرات المختلفة التي تركز عليها برامج العمل والإنتاج ، ورفع المستوى ونوعية الكوادر الهندسية وتوسيع القاعدة النظرية لمعلوماتهم ، وأن دراسة علم الإحصاء ونظرية الاحتمالات تعتبر أحد البنيات الأساسية العلمية الحديثة الهامة في حل هذه المشكلة .

ولا يسعنا في الختام إلا أن نتوجه بشكرنا وتقديرنا إلى الذين ساهموا بشكل أو بآخر في أعداد هذا الكتاب ونخص بالذكر الدكتور أقبال رسمي محمد والسيدة ناديا مال الله ، كما نتقدم بوافر الشكر والامتنان إلى الدكتورة انتصار الباجه جي لتعاونها المخلص في طباعة هذا العمل .

وأخيراً نتقدم بالشكر الخالص إلى الأخ الأستاذ الدكتور محمد إدريس فضل وإلى جميع العاملين في إدارة المركز العالي للمهن الميكانيكية - جنزور ، وإلى الأخ المهندس مصطفى حسن أبو الرقم وجميع العاملين في إدارة المعاهد والمراكز المهنية العليا بشعبية الجفارة لما أولوه من اهتمام ودعم مغوي في إظهار فكرة هذا العمل للوجود .

المؤلفون

الباب الأول

مدخل في علم الإحصاء

(Introduction to Statistics)

- 1.1 مقدمة .
- 2.1 طبيعة البيانات الإحصائية .
- 3.1 الوحدة الإحصائية والمجموعات الإحصائية .
- 4.1 المعلومات الإحصائية .
- 5.1 خطوات البحث الإحصائي .
- 6.1 تصنيف البيانات الإحصائية .
 - 1.6.1 مراجعة البيانات .
 - 2.6.1 تصميم الجداول .
 - 3.6.1 تبويب البيانات .
 - 4.6.1 جداول التوزيعات التكرارية .
 - 1.4.6.1 جداول التوزيعات التكرارية المزدوجة .
- 7.1 التمثيل والعرض البياني للبيانات .
- 8.1 التمثيل البياني للتوزيعات التكرارية .
- 9.1 تمارين .

1.1 مقدمة (Introduction)

الإحصاء (Statistics) هو العلم الذي يدرس كيفية جمع المعلومات من المجتمعات الإحصائية المختلفة سواء بالعد الشامل أو بالمعينة وكيفية تحويل هذه المعلومات إلى بيانات رقمية في جداول إحصائية ، بالإضافة إلى الأساليب المختلفة التي يمكن استخدامها لتحليل هذه البيانات تحليلًا رياضيًا لاستنتاج المقاييس المختلفة مثل المتوسط والانحراف المعياري ، وعند حساب المقاييس أو المعاملات مثل معامل الارتباط ومعامل الانحدار ومعامل الاختلاف وغيرها ، أو المؤشرات التي تدل على الاتجاهات الزمنية مثل الأرقام القياسية المختلفة ثم كيفية إجراء الاختبارات المختلفة على المقاييس والمعاملات المستنتجة من عينات للحكم على معنوياتها وتحديد أخطائها عند درجات الثقة المختلفة . وأخيرًا كيفية تفسير النتائج التي نصل إليها باستخدام هذه الأساليب في التحليل ثم توضيحها في تقرير نهائي عن موضوع الدراسة الذي أرادنا دراسته باستخدام الطريقة الإحصائية .

يتبين لنا بذلك أن علم الإحصاء يعنى بالأساليب الإحصائية التي يلجأ إليها الباحث سواء في العلوم الطبيعية أو الاجتماعية للتعرف على الحقائق الخاصة بالظواهر والمشاكل موضع البحث ، ولذلك نستطيع أن نعرفه بأنه العلم الذي يدرس إحدى طرق البحث العلمي ، حيث يوضح الخطوات التي تتبع عندما يتقرر استخدام الطريقة الإحصائية كمنهاج للبحث في أي من الأبحاث العلمية أو بمعنى آخر هو العلم الذي يدرس كيفية جمع المعلومات ثم تبويبها وتحليلها إحصائيًا ثم عرض النتائج في رسوم بيانية ومقاييس ومؤشرات ومعاملات مع التفسير الخاص بدلالة كل منها .

2.1 طبيعة البيانات الإحصائية (Nature of Statistical Data)

تقوم البيانات الإحصائية على المعالجة الرياضية للمعلومات الخاصة بعدد كبير من الوحدات الإحصائية ، أي الخاصة بمجتمعات إحصائية ضخمة عندما

تكون في شكل رقمي فقط . وبشكل عام نلاحظ أن المعلومات الخاصة بالنواحي التجارية تكون أساس معلومات رقمية مثل النقد المتداول ، المبيعات ، المخزون ، الودائع ، القروض ، رؤوس الأموال ، الكميات المنتجة ، الصادرات ، الواردات ، الدخل القومي ، أسعار الفائدة ، أسعار الصرف وغيرها .

إلا أن الكثير من المعلومات الأخرى تكون وصفية غير رقمية ، وبذلك يصبح من الضروري البدء بتحويلها إلى بيانات رقمية حتى يمكن إجراء التحليل الإحصائي الذي يتقرر إتباعه . ولاشك أن نجاح أية دراسة يتوقف إلى حد كبير على قدرة القائم على هذه الدراسة على صياغة الأسئلة التي تتضمنها الاستمارة الإحصائية التي تستخدم في جمع المعلومات الخاصة ، بحيث تساعد في الحصول على إجابات يمكن تحويلها إلى أرقام تصلح لأن تكون مادة لإجراء التحليل الإحصائي الذي يتفق وموضوع الدراسة والبحث .

فمثلا ، يمكن التعبير عن الصحة بعدد أيام المرض التي عاناها أفراد المجتمع الإحصائي موضوع الدراسة . كما يمكن التعبير عن الذكاء بإجراء اختبارات قياسية معينة . وكذلك يمكن التعبير عن آراء الناس بالنسبة لمشكلة ما بأرقام معينة . كما يمكن التعبير عن فئاتهم الوصفية بأرقام تدل عليها . ولا شك في أن هذه المشكلة تواجه الباحث بشكل خاص في أبحاث السوق حيث يتحتم عليه أن يوجه أسئلة إلى المستجوبين في المجتمع الإحصائي موضوع دراسته للحصول على إجابات وصفية ، وبذلك تظهر المهارة في كيفية صياغة هذه الأسئلة بالصيغة التي تؤدي إلى إجابات محددة يمكن تحويلها بسهولة وبدقة إلى أرقام تدل عليها وتكون في نفس الوقت قابلة للتحليل الإحصائي الذي يرغب في إجرائه . ويلاحظ في هذا الصدد أمرين ، الأمر الأول يتعلق بنتائج الدراسات والأبحاث الإحصائية التي نظرا لكونها تعتمد على القياس الرقمي والتحليل الرياضي تكون تبعا لذلك ذات مضمون

موضوعي ، إلا أن هذه النتائج من ناحية أخرى لا بد وأن تتأثر في النهاية بالتفسيرات الشخصية للقائمين بهذه الدراسات والأبحاث . فعند دراسة الارتباط بين سعر سلعة ما والطلب عليها مثلا ، قد يعتمد الدارس والباحث نموذجا رياضيا بسيطا يتضمن هذين المتغيرين فقط ، وبذلك بالرغم من أنه قد يصل إلى مقياس يدل على وجود ارتباط شديد جدا بين هذين المتغيرين ، إلا أن هذه النتيجة قد تكون مضللة بعض الشيء نظرا لإهمال المتغيرات الأخرى الكثيرة التي يمكن أن يكون لها تأثير على طلب السلعة ، مثل دخول المستهلكين وأسعار السلع الأخرى البديلة وغير ذلك من المتغيرات .

أما الأمر الثاني فيتعلق بدرجة دقة نتائج الدراسات والأبحاث الإحصائية حيث أنه بالرغم من أنها تظهر بشكل رقمي إلا أننا لا نستطيع أن نضفي عليها صفة الدقة الكاملة مثل نتائج التحليل الرياضي البحت ، ذلك لأن النتائج الإحصائية تعتمد على معلومات نحصل عليها بتوجيه أسئلة معينة إلى وحدات المجتمع موضوع الدراسة ولا نستطيع أن نجزم بأن جميع الإجابات تمثل الواقع تماما إذ لا بد أن يتسرب إليها شيء من عدم الدقة .

أن هذا الأمر ذلك لا يعني عدم جدوى نتائج الدراسات التي تجرى باستخدام المنهج الإحصائي ، حيث أن الباحث بمهارته وخبرته يستطيع أن ينقص الأخطاء التي يمكن أن تتسرب إلى المعلومات التي يبني عليها لتحليل الإحصائي إلى أدنى حد ممكن . وفي هذا المجال يلعب الوعي الإحصائي الذي يسود المجتمع العامل فيه الباحث دورا أساسيا في التأثير على مجرى الدراسات الإحصائية وعلى دقة نتائجها وأهميتها العملية تبعا لذلك . ومن ناحية أخرى يستطيع الباحث باستخدام أساليب إحصائية معينة أن يحدد مدى الخطأ في النتائج التي يصل إليها ودرجة الثقة في هذه النتائج ، خاصة عند استخدام المعاينة لجمع المعلومات التي يرغب في الحصول عليها .

3.1 الوحدة الإحصائية والمجتمعات الإحصائية

(Statistical Unit and Statistical Societies)

أن الدراسات العلمية التي تعتمد الأسلوب الإحصائي منهاجا للبحث تبدأ عادة بتحديد المجتمع الإحصائي موضوع الدراسة وبالتالي الوحدة الإحصائية التي تكون هذا المجتمع . ففي دراسة عن الصناعة مثلا ، يجب تحديد المؤسسات الصناعية التي تتضمنها الدراسة سواء من ناحية نوع نشاطها أو من ناحية حجمها قياسا بعدد العاملين فيها أو برؤوس الأموال الموظفة فيها وغيرها . كذلك عند دراسة مثلا ، ميزانية الأسرة فإنه يجب تحديد الأسر التي تتضمنها الدراسة سواء من ناحية سكنها في الريف أو في المدن أو فيهما سوية ، وكذلك من ناحية مستوى دخلها وغيرها .

بذلك يتبين لنا أن المجتمع الإحصائي هو مجموع الوحدات مهما كان نوعها ، أفراد أو أسر أو مؤسسات أو مساكن أو مدارس والتي تكون موضوع الدراسة التي يرغب الباحث في القيام بها . ولأنك أن تحديد أي مجتمع إحصائي يستلزم حتما تعريف دقيق للوحدة الإحصائية التي يتكون منها هذا المجتمع . فإذا كنا بصدد دراسة المؤسسات الصناعية يجب قبل البدء بجمع المعلومات المطلوبة تحديد ما هو المقصود بمؤسسة صناعية ، وكذلك توضيح المشاكل التي يمكن أن نواجهها عند التعرف على المؤسسات الصناعية ميدانيا ، مثل مؤسسة تعمل في الزراعة والصناعة سويا ، أو مؤسسة لها فروع في أقسام إدارية مختلفة في الدولة موضوع الدراسة ، أو مؤسسة تعمل في نشاطات مختلفة ليست جميعا موضع البحث ، ووضع الحلول التي يجب أن يتبعها القانون عند مواجهة أي من هذه المشاكل .

وحتى إذا كان بالإمكان جمع المعلومات الميدانية على أسس موحدة فإنه لا يكفي للتحديد الوحدة الإحصائية والمجتمع الإحصائي الذي تستهدفه الدراسة ، بل يجب كذلك تعريف كل المصطلحات التي تتضمنها أسئلة البحث

تعريفا واضحا تماما ولا شك أن عجز الباحث عن إعداد هذه التعاريف الواضحة غير المبهمة يجعل تلك المعلومات التي تجمع من وحدات المجتمع الإحصائي موضوع الدراسة خاضعة للتفسيرات الشخصية من قبل المستجوبين أو من قبل القائمين بالعمل ميدانيا وبذلك لا يمكن المقارنة بينهما .

كذلك لا يمكن تصنيفها وتبويبها على أسس موحدة ، فإذا أُجري مثل هذا التصنيف والتبويب على معلومات من هذا النوع يكون التحليل الإحصائي الذي يجري عليها فيما بعد عملا غير دقيق حيث يؤدي إلى نتائج مضللة وذلك لأن الإجابات التي نحصل فيها على أسئلة مبهمة غير واضحة تكون في الواقع إجابات لما يفهمه المستجوبون وهو بدون شك فهم يختلف من مستجوب لآخر نتيجة هذا الإبهام وعدم الوضوح .

ويجب الإشارة هنا إلى أن كثير من الكلمات قد تكون واضحة تماما في ذهن الباحث ولكنها تفهم بصورة مختلفة من قبل أولئك الأشخاص المستجوبين على الرغم من بساطة هذه الكلمات ، والأمثلة على ذلك كثيرة مثل الحالة الزوجية ، الجنسية ، المستوى التعليمي ، المهنة ، النشاط الاقتصادي ، إنتاج المؤسسة ، المشتغلون في المؤسسة وغيرها .

لذلك يجب التأكيد على الأهمية القصوى لتعريف جميع الألفاظ التي ترد في الأسئلة مهما كانت تبدو لنا واضحة وبسيطة ، وذلك لأن معنوية النتائج التي نصل إليها من الدراسات الإحصائية تتوقف أساسا على التعاريف الواضحة غير المبهمة لوحدات المجتمع الإحصائي موضوع أي من هذه الدراسات ولجميع المصطلحات والألفاظ التي تتضمنها الأسئلة التي تستخدم في جمع المعلومات الخاصة بهذه الدراسات .

4.1 المعلومات الإحصائية (Statistical Information)

تسمى المعلومات التي تجمع خصيصاً لدراسة إحصائية معينة بالمعلومات الأولية (Preliminary Information) ، وذلك مثل المعلومات التي تجمع في تعداد السكان أو في التعداد الصناعي أو في استقصاء ميزانية الأسرة وغيرها. والميزة الأساسية لمثل هذه المعلومات هي إمكانية التحكم في صياغة الأسئلة التي تتضمنها استمارة التعداد أو الاستقصاء وتوضيح كل ما تتضمنه هذه الأسئلة من ألفاظ وتعابير بحيث يمكن تجنب سوء الفهم وتقليل الأخطاء التي تترتب على ذلك إلى حد كبير .

إلا أن هناك الكثير من المعلومات التي تجمع لأغراض أخرى غير إحصائية سواء في مختلف الإدارات الحكومية أو في مختلف المؤسسات الاقتصادية الخاصة والتي يمكن أن تظهر في سجلات هذه الإدارات وهذه المؤسسات ، وتكون بذلك مصدراً لكثير من الإحصاءات . هذه المعلومات تسمى بالمعلومات الثانوية (Secondary Information) ، حيث أن استخدامها لأغراض إحصائية يكون مجرد استخدام ثانوي ، فهي تنظم أساساً لأغراض إدارية أو قانونية . والأمثلة على ذلك كثيرة منها سجلات المواليد والوفيات والزواج والطلاق ، سجلات الجمارك ، سجلات الشرطة والقضاء ، السجلات المحاسبية للمؤسسات الاقتصادية المختلفة وغيرها .

وعند استخدام هذه المعلومات كمصادر للإحصاءات الخاصة بها يجب أن يكون ذلك بعناية فائقة ، حيث أن هذه المعلومات قد لا تكون شاملة شمولاً كاملاً لجميع وحدات المجتمع الإحصائي موضوع الدراسة ، كما قد لا تكون مسجلة بالأسلوب الصحيح الذي يتفق مع اعتبارها مصدراً إحصائياً جيداً . ولذلك قد يتعين علينا إجراء بعض التعديلات على المعلومات من هذا

النوع حتى تصبح صالحة لأن تكون مادة أولية جيدة للدراسة الإحصائية المطلوب القيام بها . مثلاً ، لمعرفة التغيرات الموسمية في مبيعات إحدى المؤسسات أو في النقد المتداول يكون من الواجب إعداد البيانات الإحصائية عن المبيعات أو النقود المتداولة في فترات شهرية أو ربع سنوية بدلاً من إعدادها على أساس سنوي .

لذلك يستحسن تنظيم السجلات المستخدمة في مختلف الإدارات الحكومية وفي مختلف المؤسسات الاقتصادية بالتعاون بين القائمين بالعمل فيها وبين المسؤولين عن إعداد الإحصاءات حتى يمكن الاتفاق مقدماً على طرق الحصول على البيانات التي تتضمنها هذه السجلات ، وعلى الأسئلة التي توجه وعلى تعريف الألفاظ والتعابير التي ترد في هذه الأسئلة ، وعلى جميع الصعوبات التي يمكن أن تظهر عند العمل وكيفية معالجة تلك الصعوبات .

لذا يتعين علينا قبل استخدام أي إحصاءات في أي من الأبحاث العلمية ومهما كان مصدر هذه الإحصاءات أولياً أو ثانوياً ، إن نتأكد من كفاءة الحصول عليها ، ومن درجة شمولها للمجتمع الإحصائي موضوع البحث ، ومن الفترة الزمنية التي تتعلق بها ، ومن التعاريف المختلفة التي استخدمت بالنسبة للألفاظ والتعابير الخاصة بموضوعها ، وإن أي تهاون في التأكد من هذه القواعد الأساسية التي يقوم عليها أي إحصاء مهما كان مصدره أولياً أو ثانوياً يمكن أن يؤدي إلى نتائج مضللة ، وبذلك يكون البحث عديم الجدوى ، بل مؤدياً إلى أخطاء فادحة في مسيرة الأعمال التي تعتمد على نتائجه .

5.1 خطوات البحث الإحصائي (Statistical Research Steps)

يقصد بالبحث الإحصائي دراسة أي موضوع ما أو مشكلة ما في أي من المجالات العلمية ، الاجتماعية أو الطبيعية ، باستخدام الطريقة الإحصائية منهاجاً للبحث وأداة للتوصل إلى إجابات عن الأسئلة المختلفة

التي يطرحها الباحث في دراسته وإلى التأكد من الفرضيات النظرية التي بدأ بها بحثه لقبولها إذا تبين من البحث معنوياتها أو استبعادها إذا تبين عدم معنوياتها . ولإجراء مثل هذا البحث تجرى دراسة استقصائية ميدانية لجمع المعلومات التي يرى الباحث ضرورتها . وعند إجراء مثل هذه الدراسة تتبع الخطوات التالية :

1- تحديد موضوع البحث تحديداً واضحاً من جميع النواحي بحيث يكون مفهوماً مقدماً المشكلة التي يرغب الباحث في دراستها أو الفرضية النظرية التي يرغب في اختبارها والحكم عليها . فإذا كنا بصدد دراسة الأجور في المؤسسات الصناعية ، مثلاً يكون من الواجب أن نحدد مقدماً هل نحن بصدد دراسة معدلات الأجور أو كسب العمل . وهل تشمل الدراسة جميع المشتغلين أو بعض الفئات منهم وما هي هذه الفئات . وهل تتضمن الدراسة المكافآت الإضافية التي حصل عليها المشتغلون أو تقتصر على الأجور الأساسية فقط . ومن الواضح أن تحديد موضوع البحث تحديداً شاملاً يرشد الباحث في تحديد المعلومات المختلفة التي يسأل عنها ويحاول الوصول إليها .

2 - إذا لم يكن من الممكن جمع المعلومات من جميع وحدات المجتمع الإحصائي الذي يشملته البحث بسبب ضخامة التكاليف ، واتساع الجهاز الإداري والفني الذي يحتاجه العد الشامل ، وقد يكون إعداد مثل هذا الجهاز أمراً متعذراً بالإضافة إلى طول الوقت الذي يمضي قبل الوصول إلى نتائج الدراسة ، وكثرة الأخطاء التي يمكن أن تترتب على العمل على نطاق واسع ، لذلك نجمع المعلومات من عينة من الوحدات على أن تتم معالجة نتائج العينة معالجة رياضية معينة ، بحيث يمكن التوصل إلى تقدير المقاييس والمعاملات الخاصة بالمجتمع الذي تمثله العينة موضوع الدراسة وحساب أخطاء هذه التقديرات ودرجة الثقة فيها .

في هذه الحالة لا بد أن يتقرر مقدماً نوع العينة وحجمها وطريقة سحبها من المجتمع المستهدف في الدراسة . ولا شك في أن مثل هذه القرارات تتوقف على مدى تجانس وحدات المجتمع ودرجة الدقة التي يقبل بها الباحث ونوع المعلومات التي سوف يسأل عنها والمبالغ المخصصة للدراسة .

3 - إعداد استمارة البحث ، وهي عملية شاقة حيث تحتاج إلى خبرة وفهم دقيق لموضوع الدراسة بالإضافة إلى الخبرة الإحصائية والممارسة الطويلة لهذا النوع من الأعمال ، إذ لا بد أن تأتي الأسئلة واضحة ومحددة من حيث ألفاظها وصياغتها فلا تتعرض إلى أي نوع من الغموض والتأويل . كما لا بد أن تأتي الأسئلة في صورة تؤدي إلى إجابات قابلة للمعالجة الإحصائية ، أي قابلة للتبويب ضمن أرقام في جداول يمكن أن يجري عليها التحليل الإحصائي الرياضي . وبالرغم من ضرورة الخبرة بهذا النوع من الأعمال ، إلا أن تجربة الاستمارة ميدانياً أمر حتمي للتأكد قبل استخدامها من أن فهم المستجوبين لأسئلتها يطابق تماماً فهم الباحثين لها ، ولتعديل أي لفظ أو تعبير يتبين من التجربة سوء الفهم له .

وبالرغم من أن الكثير من المعلومات الخاصة بالمؤسسات الاقتصادية تتوفر في سجلات هذه المؤسسات ، مثل المعلومات الخاصة بالتكاليف والمشتغلين والمبيعات ، إلا أن الحاجة تظهر أحياناً لإجراء دراسات استقصائية ميدانية للتعرف على أمور معينة تتعلق بأعمالها ، وذلك مثل أبحاث السوق .

4 - جمع المعلومات ويكون ذلك بإحدى الطرق الآتية :

a) إرسال الاستمارات إلى المستجوبين بالبريد ومطالبتهم بإجابة أسئلتها وفق التعليمات المرفقة ثم إعادتها إلى الدائرة القائمة بالدراسة في ملف مرفق عليه العنوان الخاص بهذه الإدارة. ولا شك أن نجاح هذه الطريقة يتوقف إلى حد

كبير على الوعي الإحصائي للمستجوبين ، حيث أن هذا الوعي يجعلهم يدركون أهمية المعلومات التي تطلب منهم ولذلك يستجيبون بإعطاء الإجابة الدقيقة .

(b) إرسال فريق من الموظفين الذين تم تدريبهم على مقابلة المستجوبين وكيفية إقناعهم بإعطاء المعلومات الدقيقة وكيفية طرح الأسئلة وتدوين إجاباتها وأخيراً كيفية مراجعة الاستثمارات بعد تعبئتها للتحقق من عدم وجود أي خطأ وأن جميع الأسئلة قد أجيب . ومن الواضح أن العمل بهذه الطريقة يحتاج إلى تنظيم وأشراف دقيق للتأكد من قيام الموظفين بجمع المعلومات ولمساعدتهم في مواجهة المشاكل التي قد تظهر أثناء العمل ولتدقيق الاستثمارات بعد تعبئتها حتى يمكن تصحيح أي خطأ يظهر فيها وهي لا تزال قيد التداول ميدانياً .

(c) إرسال فريق من الموظفين إلى أماكن معينة لملاحظة ما يحدث وتدوين هذه الملاحظات على بطاقات معدة لذلك الغرض. وتتبع هذه الطريقة عند إجراء دراسات عن حركة المرور أو عن سلوك الأطفال أثناء لعبهم.

(d) توجيه أسئلة محدودة بواسطة جهاز الهاتف، إلا أن هذه الطريقة تكون في الغالب متحيزة حيث أن الأسر التي لديها أجهزة هواتف لا يمكن اعتبارها عينة تمثل المجتمع موضوع البحث تمثيلاً صادقاً. ويمكن اتباع هذه الطريقة في بعض الدراسات الخاصة بآراء الناس بالنسبة لبرامج المذياع أو التلفاز .

(e) إجراء التجارب أو القياسات المحددة ، وتتبع هذه الطريقة عند مراقبة الإنتاج كماً وكيفاً للتأكد من سير العمل وفق المواصفات والمعدلات المحددة بالنسبة للماكينات أو للعمال ، أو عند إجراء التجارب الزراعية وتسجيل نتائجها التي تكون بعد ذلك موضع التحليل الإحصائي .

5 - المراجعة الأخيرة للمعلومات المدونة في الاستمارات للتأكد من تماسكها، أي عدم تناقضها فيما بينها أو عدم تناقضها مع ما متوقع لها . ولغرض الكشف عن السهو أي عن الأسئلة التي ليس هناك أي إجابة لها . بعد ذلك تجرى بعض العمليات الحسابية لكي يستنتج من واقع المعلومات المعطاة في كل استمارة مقياساً مطلوباً ، مثل حساب القيمة الصافية المضافة لكل مؤسسة صناعية من واقع المعلومات التي أعطتها المؤسسة عن مبيعاتها خلال عام معين والمخزون المتوفر لديها في أوائل وأواخر العام المعني ، وكذلك مشترياتها من المواد الخام والوقود خلال العام والمخزون من هذه المواد في أول العام وفي آخره .

بعد ذلك تبدأ العمليات الخاصة بالتبويب الآلي وهي ترميز المعلومات غير الرقمية ، ثم تقليب البطاقات وبعد ذلك تصنيفها وفق الفئات المطلوبة ، ثم التبويب الآلي أي وضع البيانات بعد تصنيفها في جداول باستخدام الآلات الخاصة بذلك . ونشير هنا إلى أن جميع العمليات بعد الترميز تجري آلياً ويراجع إجراءاتها آلياً كذلك .

6 - تحليل البيانات الإحصائية أي المعلومات بعد تصنيفها وتبويبها ، وتفسير النتائج التي تترتب على هذا التحليل ، ومن ثم إعداد تقرير عن الدراسة معزراً بالجدول والرسوم البيانية والمقاييس والمعاملات والمؤشرات التي أمكن التوصل إليها ، مع ملاحظة أن يتضمن التقرير تعريف موجز بجميع الخطوات التي أتبعته وتوضيح أسباب وخلفيات أي قرار أتخذ بصدد هذه الخطوات ، مثلاً لماذا تقرر استخدام هذا النوع من العينات ولم يتبع أي من العينات الأخرى ، وعلى أي أساس حدد حجم العينة بعدد معين من المفردات وغيرها .

6.1 تصنيف البيانات وتبويبها

(Classification and Tabulation of Data)

1.6.1 مراجعة الساتات (Revision of Data)

قبل البدء في تصنيف البيانات ووضعها في جداول مناسبة يتم عادةً مراجعتها بهدف اكتشاف بعض الاستمارات التي قد تحتوي بيانات متناقضة أو نقص في الإجابات ، والتي يجب إعادتها إلى الميدان لاستيفائها أو إلغاؤها في حالة عدم التمكن من تصحيحها . فمثلاً إذا كان البيان المطلوب هو عن أجور بعض العمال الصناعيين فيجب أن يكون أساس الأجر موحداً حيث أن بعض المصانع تعطي أجراً يومياً والأخرى أسبوعياً أو شهرياً ، فيجب إرجاع جميع الأجور إلى أساس موحد أسبوعي مثلاً حتى تكون جميع الوحدات المستخدمة في البحث متجانسة ومن نوع واحد .

2.6.1 تصميم الجداول (Design of Tables)

بعد الانتهاء من مراجعة البيانات ، يتم إعادة تنظيمها بطريقة تسهل من دراستها وذلك للاستفادة منها على نحو أفضل . من ذلك يتم تصنيفها أي تقسيمها إلى مجموعات متجانسة وجدولتها أي وضعها بصورة جداول تلخيصية . وعادةً يتوقف هذا التقسيم على طبيعة البيانات وعلى الغرض الذي نسعى إليه من عمل البحث والدراسة . وبشكل عام لا توجد طريقة موحدة واحدة لعمل هذه الجداول إلا أن هناك قواعد عامة يجب مراعاتها وأخذها بعين الاعتبار عند تصميم الجداول ومنها :

(a) أن يكون عنوان الجدول المقصود واضحاً ومختصراً ومحدداً لما يحتويه من معلومات .

(b) أن تكون عدوين الصفوف والأعمدة مختصرة وواضحة وضوحاً تاماً .

(c) أن ترتب البيانات في الجداول وفق تسلسلها الزمني أو حسب أهميتها من الناحية الوصفية .

(d) يجب ترقيم الصفوف أو الأعمدة لتسهيل الإشارة إلى بيانات الجداول .

(e) أن يوضح المصدر الذي أخذت منه بيانات الجداول .

(f) أن توضع وحدات القياس المستخدمة بدقة .

3.6.1 تبويب البيانات (Grouping of Data)

ويقصد بالتبويب في علم الإحصاء وضع البيانات الإحصائية على شكل جداول (Tables) ، بحيث تمكننا من عرضها بصورة تلخص معالمها، وتساعد على استخلاص النتائج منها . ويتم عادةً تبويب البيانات على أساس تقويم زمني أو نوعي أو كمي أو جغرافي . أو على أساس خليط من هذه الأسس المختلفة التقسيم . فمثلاً الجدول (1-1) يبين قيمة المنتجات الصناعية في إحدى الدول العربية مبوبة على أساس أحد الأصناف السابقة وهو التبويب الزمني ، حيث يتم العرض الخاص بكل وحدة زمنية وفي هذا المثال هي السنة .

وهكذا يتم تبويب البيانات على أساس الأسس الأخرى المذكورة سابقاً . وقد يشمل التبويب على أكثر من واحد من هذه الأسس ، فقد يكون التبويب نوعياً وكمياً مثلاً في نفس الوقت وذلك إذا تم تصنيف المنتجات حسب حجم المصنع المنتج لها وذلك بالنسبة لكل نوع من أنواع الصناعة .

جدول (1-1)

قيمة المنتجات الصناعية في إحدى الدول
العربية خلال السنوات (1980 - 1985)

السنة	قيمة المنتجات (بالآلاف الديناري)
1980	482359
1981	519697
1982	538723
1983	622451
1984	734523
1985	754291

4.6.1 جداول التوزيعات التكرارية

(Frequency Distribution Tables)

في كثير من الأحيان نلاحظ أنه عند الحصول على بيانات حول ظواهر أو تجارب معينة تكون هذه البيانات في وضع عشوائي أي غير خاضعة لأي نوع من الترتيب أو التصنيف ، وعادة ما تكون البيانات في مثل هذه الحالات كثيرة العدد ، وفي هذه الحالات يصعب فهم وتوضيح أي معالم أو صفات لهذه البيانات بسهولة . ولكن عند وضعها في صورة منظمة يمكن بها عندئذ توضيح معالمها الأساسية بسهولة ، فمثلاً لو كان لدينا درجات 30 طالباً في مادة ما مبينة على النحو التالي :

21 , 32 , 54 , 75 , 43 , 75 , 86 , 54 , 86 , 32 , 24 , 32 , 63 , 75
96 , 43 , 51 , 21 , 63 , 96 , 21 , 86 , 46 , 32 , 34 , 54 , 63 , 51
67 , 21.

من الواضح وبالرغم من أن البيانات قليلة العدد إلا أن هناك عدة معلومات مبعثرة قد يصعب استنتاجها وتوضيحها من هذه الصورة للبيانات ببساطة ، ولكن وبمجهود بسيط نستطيع الحصول على معلومات عديدة توضح المعالم الأساسية الخاصة بهذه البيانات . فمثلاً إذا تم ترتيب هذه البيانات ترتيباً تصاعدياً أو ترتيباً تنازلياً مع توضيح تكرار كل قيمة ، لتمكننا من توضيح هذه البيانات على النحو التالي :

جدول (1 - 2)

المجموع	96	86	75	67	63	54	51	46	43	32	24	21	الدرجة
30	2	3	3	1	3	3	2	1	2	4	1	4	التكرار

بهذه الطريقة أصبح من الواضح معرفة معلومات عديدة كانت غير واضحة في الوضع السابق للبيانات ، فمثلاً من السهل الآن معرفة أكبر درجة وأصغر درجة كما يمكن معرفة عدد الناجحين ببساطة ، أي بهذه الطريقة استطعنا توضيح المعالم الأساسية لهذه البيانات وإعطاء فكرة عامة عن درجات الطلبة . ومن الواضح أنه لو كانت الدرجات كثيرة ومنتشرة داخل مجال متسع ووضعت على الشكل السابق والذي يبين درجات الطلبة مفردة مع تكرار كل درجة فسوف نجد أن الجدول (1-2) لن يفي بالغرض المطلوب.

من هذه الطريقة وهي تلخيص ووصف البيانات بطريقة أبسط توضح المعالم الرئيسية أي وضع البيانات في صورة مختصرة ومنظمة تعطي فكرة عامة عن هذه البيانات ، حيث نجد أن الجدول الذي نحصل عليه لن يختلف كثيراً عن وضع البيانات الأصلي لذلك فقد وجدت طريقة أخرى أكثر اختصاراً وأشمل فائدة من الوضع الأصلي للبيانات يمكن بواسطتها وضع البيانات في جدول يبين ويوضح الخصائص العامة لهذه البيانات .

وتتلخص هذه الطريقة في الواقع ، في أنه بدلاً من التعامل مع البيانات بالطريقة السابقة فإن المفردات تقسم إلى مجموعات أو فئات متجانسة بحيث تشمل كل مجموعة أو فئة عدداً من القيم المتقاربة من بعضها ، بحيث لا تنتمي كل مفردة من المفردات إلا إلى واحدة فقط من هذه الفئات (Sets or Classes) .

ويتوقف عدد هذه المجموعات أو الفترات (Intervals) على المدى (Range) بين أكبر وأصغر قيمة من قيم المفردات التي لدينا حيث يقسم المدى إلى عدد مناسب من الفترات أو الفئات لكي تضم كل فئة من الفئات مجموعة من القيم المتقاربة . إن تحديد أطوال هذه الفئات وعددها يتوقف عادة على طبيعة البحث ودرجة البحث ودرجة التلخيص المطلوبة .

وقد تكون الفترات أو الفئات متساوية الطول أو غير متساوية وذلك حسب طبيعة البيانات والغرض المعمول من أجله الجدول ، وبعد ذلك تعامل البيانات كمجموعات جديدة وليس كأعداد منفصلة كما كانت عليه قبل التجميع والتجميع ، وبالرغم من أنه لا توجد طريقة وحيدة لوضع جداول التوزيعات التكرارية إلا أنه هناك بعض النقاط الأساسية والخطوات الرئيسية التي يجب مراعاتها حتى تكون المعلومات التي نحصل عليها من الجداول أقرب إلى واقع البيانات وأكثر دقة ومن أهم هذه الخطوات والنقاط :

1. نحدد المدى (Range) الذي تنتشر فيه البيانات ويمثل القيمة الكبرى للبيانات أي القصوى (Maximum) مطروحاً منه القيمة الصغرى للبيانات أي (Minimum) أي أن المدى يساوي :

$$R = \text{Maximum (Max.)} - \text{Minimum (Min.)} \dots\dots\dots (1-1)$$

2. نقسم المدى إلى فترات أو فئات متساوية الطول بحيث يكون عددها مناسباً. وقد أعتبر العدد من 5 إلى 25 عدداً مناسباً للفترات. وقد وضع العالم (استيرجس) علاقة يمكن الاستعانة بها لتحديد عدد الفترات وهي :

$$(2 - 1) \dots\dots\dots \text{عدد الفترات} = 1 + 3.3 \text{ Log } n$$

حيث إن :

n - هو عدد البيانات . وليس من الضروري إتباع هذه العلاقة في تحديد عدد الفترات ولكن الأمر متروك للباحث وخبرته في وضع الجداول .

وهنا يجب الإشارة إلى أن عدد الفترات يجب أن يكون مناسباً وليس بالعدد الكبير حتى تكون معظم البيانات مشتتة بين فترات عديدة ، وبذلك لن يكون هناك فرق كبير بين البيانات بوضعها الأصلي وبين الجداول ولن يوضح الجدول المعالم الأساسية التي وضع من أجلها ، كذلك يجب أن لا يكون عدد الفترات صغير جداً حتى لا نضطر إلى دمج معظم المعلومات معاً وفي هذه الحالة لن نستطيع توضيح وإعطاء فكرة عامة عن البيانات .

3. تحديد الحدود العليا والدنيا الفعلية للفئات حتى لا يكون هناك فجوات أو تداخل ما بين هذه الفئات .

4. يكون اختيار طول الفئة من الأعداد التي يمكن التعامل معها بسهولة .

ويجب الإشارة إلى أنه عند تبويب جداول التوزيعات التكرارية فإن البيانات الأصلية تفقد في الواقع ، ولا يمكن الرجوع من الجداول إلى تلك البيانات الأولية ولكن مقابل ذلك يتم الحصول من الجداول على المعلومات

والاستنتاجات التي توضح المعالم الأساسية للبيانات بسهولة ، لذا يجب عند تكوين الجداول إتباع الخطوات الأساسية التي تم التعرض لها حتى تكون المعلومات التي سنحصل عليها من الجداول قريبة من الواقع الأولي لهذه البيانات .

وقد ذكرنا أيضاً أن طول الفترات في الجداول يجب أن يكون متساوياً وذلك لسهولة التعامل معها ، ولكن في بعض الحالات قد تستخدم فترات غير متساوية الطول لوجود غرض معين من وراء ذلك . فمثلاً إذا كان الغرض من الدراسة الاهتمام ببعض الفترات والتركيز عليها مع عدم الاهتمام لباقي الفترات الأخرى ، فإنه يتم دمج الفترات التي لا تهم الباحث في فترة واحدة ويكون الجدول في هذه الحالة غير متساوي الفترات . كذلك إذا كان التكرار (Frequency) لبعض الفترات صغيراً جداً مقارنة بباقي الفترات يمكن وضع هذه الفترات معاً .

كذلك هناك نوع آخر من جداول التوزيعات التكرارية تجدر بنا الإشارة إليه وهي الجداول ذات الفترات المفتوحة (Open Intervals) ، حيث يصادفنا أن يكون الحد الأدنى للفترة الأولى أو الحد الأعلى للفترة الأخيرة غير محدد ، مثل أن نقول الفترة الأولى أقل من 100 أو الفترة الأخيرة من 400 فأكثر . وهذا النوع من الجداول عادةً ما يكون قليل الأهمية حيث إن مثل هذا النوع لا يمكننا منه حساب مقاييس إحصائية هامة يجب إيجادها كما سيتم شرحه في الأبواب القادمة من هذا الكتاب . وبالرغم من ذلك نجد أن لهذا النوع من الجداول أي الجداول المفتوحة بعض الاستخدامات في بعض الحالات .

ولتوضيح طريقة تكوين جداول التوزيعات التكرارية سنقوم بدراسة المثال التالي والذي سيوضح الخطوات الأساسية التي يجب مراعاتها عند تبويب البيانات في جداول .

مثال (1-1)

أرادت شركة صناعية كبرى متخصصة بصناعة الحاسوب ورفائق الكمبيوتر المختلفة أن تدرس كفاءة المهندسين العاملين لديها للتعرف على مدى ملائمتهم لإعمالهم الحالية واختيارهم لتمكان أنمناسب صمم مجاز عملهم . لأجل ذلك تم اختيار عينة عشوائية (Random Sample) من (50) مهندس من بين المهندسين العاملين بها والبالغ عددهم (1500) مهندس حيث أجري لهم اختبارين ، الاختبار الأول يقيس درجة نكاء المهندس وقدرته على التصرف في مواقف معينة ، والاختبار الثاني يقيس درجة المهارة اليدوية في سرعة الحركة . فكانت البيانات للدرجات التي حصل عليها المهندسون الخمسون كما مبين في الجدول (1- 3) .

جدول (1 - 3)

العينة العشوائية للخمسين مهندس

الرقم	الرقم العشوائي للمهندس في العينة	الدرجة في اختبار النكاء	الدرجة في اختبار المهارة اليدوية
1	0624	110	74
2	1434	123	68
3	0753	109	59
4	1101	104	46
5	1481	111	82
6	0962	131	46
7	0416	132	70
8	0045	127	50
9	0629	91	64
10	0483	126	49
11	1121	135	87
12	0085	116	59
13	0817	111	61
14	1383	113	59
15	1129	119	89

56	114	0612	16
48	118	1290	17
69	101	0849	18
56	118	1473	19
73	119	1082	20
56	102	0895	21
63	106	0302	22
48	124	0797	23
44	101	1311	24
57	97	0942	25
47	95	1499	26
68	121	0097	27
58	105	0009	28
70	107	0557	29
66	115	0380	30
55	110	0262	31
72	115	0589	32
54	101	0018	33
56	118	1327	34
70	121	0318	35
53	122	1337	36
75	128	0717	37
53	144	1299	38
58	119	0022	39
52	121	0609	40
77	101	0939	41
72	107	0778	42
61	141	1288	43
46	120	0089	44
69	102	0203	45
68	121	0967	46
63	133	1355	47
69	107	1023	48
60	103	0067	49
61	119	0581	50

لتوضيح كيفية تلخيص البيانات الخاصة بدرجات الذكاء للمهندسين مثلاً
في جدول توزيع تكراري ، نتبع الخطوات والمراحل التالية التي توضح لنا
كيفية عمل ذلك :

1. نقوم باختيار الفئات وكما شرحنا سابقاً مع أنه لا توجد قاعدة ثابتة لتحديد
عدد وأطوال الفئات ، حيث إن ذلك يعتمد على طبيعة البيانات المستخدمة
وعلى مستوى الدقة المطلوبة . عند النظر إلى بيانات ومفردات
الجدول (3-1) ، نجد المدى لهذه البيانات حيث إن أكبر درجة في اختبار
الذكاء هي 144 وأقل درجة هي 91 إذاً المدى يساوي :

$$R = 144 - 91 = 53$$

فلو قسّمنا هذا المدى إلى 11 فئة نجد أن طول الفئة سيكون 5 درجات ، وإذا
قسّمناه إلى 6 فئات يكون طول الفئة 10 ... وهكذا .

2. بعد اختيارنا لعدد الفئات وأطوالها يجب علينا توضيح حدود هذه الفئات
بحيث لا تتداخل مع بعضها البعض ، فمثلاً لو قسّمنا الدرجات إلى 11 فئة
طول كل منها 5 درجات فقد يتسرع أحدنا ويخطئ بكتابة الفئات على الصورة
التالية :

$$\begin{array}{c} 90 - 95 \\ 95 - 100 \\ 100 - 105 \\ | \\ | \\ | \\ 140 - 145 \end{array}$$

والخطأ هنا واضح في كتابة هذه الفئات حيث لا يمكن معرفة فيما إذا كان المهندس الذي حصل على 95 درجة ينتمي إلى الفئة الأولى أو الفئة الثانية ، والمهندس الذي حصل على 105 درجة ينتمي إلى الفئة الثالثة أو الرابعة وهكذا . وللتغلب على ذلك يمكن كتابة الفئات كما يلي :

$90 - 94$
 $95 - 99$
 $100 - 104$
 \vdots
 $140 - 144$

إن توزيع حدود الفئات بهذه الصورة يكون صحيحاً فقط إذا كان المتغير الذي ندرسه متغيراً غير مستمر (Discrete Variable) ، أي لا يأخذ قيمة كسرية أما إذا كان المتغير الذي ندرسه مستمراً (Continuous Variable) فإنه يكون من الخطأ كتابة حدود الفئات بهذا الشكل حيث يجب ألا نترك أي ثغرات بين الفئات .

فالمهندس الذي حصل على 99.5 درجة لا نعرف ما إذا كان ينتمي إلى الفئة الثانية أو الثالثة ، لذلك في هذه الحالة ، يجب توضيح جميع حدود الفئات وكتابتها بصورة واضحة وصالحة لدراسة جميع المتغيرات المستمرة وذلك بأن نجعل كل فئة تبدأ مباشرة حيث تنتهي الفئة السابقة لها دون أن يحدث تداخل بين الفئات ودون أن نترك أية ثغرات بينها كما يلي :

$$\begin{array}{r}
 95 - 90 \\
 100 - 95 \\
 105 - 100 \\
 \vdots \\
 \downarrow \\
 140 \text{ وأقل من } 145
 \end{array}$$

وللاختصار يمكن أن نحدد فقط بداية الفئة ونترك نهايتها لتحديد ضمناً من الفئة التالية لها ، وفي هذه الحالة ينبغي أن نحدد نهاية الفئة الأخيرة على الشكل التالي :

$$\begin{array}{r}
 - 90 \\
 - 95 \\
 - 100 \\
 \vdots \\
 \downarrow \\
 140 \text{ وأقل من } 145
 \end{array}$$

ومعنى هذا أن الفئة الأولى تشمل كل القيم التي تساوي 90 أو تزيد عنها بحيث نقل عن 95 وعليه فإن القيمة 95 تدخل ضمن الفئة الثانية والقيمة 140 ضمن الفئة الأخيرة وهكذا .

3. بعد اختيار الفئات وتحديد حدودها نكون جدولاً (Table) ، يحتوي على عدد من الصفوف مساوي لعدد الفئات ثم نقوم بتفريغ البيانات الأولية ، مفردة بعد الأخرى كل في الفئة التي تنتمي إليها وذلك بوضع علامة (/) أمام هذه الفئة ونستمر في عملية التبويب هذه حتى ننتهي من تفريغ جميع

مفردات البيانات الأولية وذلك بوضع عدد من العلامات مساوٍ لعدد المفردات الأولية التي لدينا ويلاحظ عادة أثناء التفريغ أننا نقوم بوضع كل أربعة علامات بجوار بعضها على الشكل التالي (////) أما العلامة الخامسة فتشطب الأربعة السابقة وتكون حزمة على الشكل التالي (////) والغرض من ذلك هو تسهيل عد هذه العلامة التي تم رصدها أمام كل فئة (Class) ، والذي سوف نطلق عليه اسم التكرار أو ما يعرف (Frequency) .

أن الجدول (4-1) ، يبين التبويب الخاص بدرجات الذكاء للمهندسين الخمسين في العينة العشوائية للمثال السابق . ويلاحظ في هذا الجدول أن الفئات تكتب في العمود الأول ثم أخذت البيانات الأولية الخاصة بدرجات الذكاء مفردة بعد الأخرى ، وقد وضعت علامة لكل مفردة أمام الفئة التي تنتمي إليها هذه المفردة . وبعد الانتهاء من عملية التفريغ نسجل عدد العلامات في العمود الأخير من الجدول والذي يمثل تكرار الفئة (f_i) .

إن مجموع التكرارات يجب أن يساوي عدد المفردات التي تم تفريغها . إن الجداول على شكل الجدول (4-1) ، والتي تتكون من الفئات والتكرارات تسمى بجداول التوزيعات التكرارية (Frequency Distribution Tables) .

وكما أشرنا سابقاً أنه عند تكوين الجداول التكرارية تضيع معالم القيم الأولية الأصلية ، ولا نعرف شيئاً عن أي مفردة من المفردات الأصلية حيث أنها سوف تنتهي إلى فئة معينة محددة بحدين معلومين مثل 100 وأقل من 105 . ولكن وبفضل هذه الجداول نستطيع معرفة الدرجة التي حصل عليها أي من المهندسين الثمانية المنتمين إلى هذه الفئة ، وما إذا كانت هذه الدرجة تقع بالقرب من بداية الفئة 100 أو بالقرب من نهايتها 105، لذلك نفترض أن جميع المهندسين في كل فئة حصلوا على درجات متساوية وأن كل منها يساوي مركز هذه الفئة أي منتصف المدى بين الحدين الأدنى والأعلى للفئة .

فمثلاً مركز الفئة 100 وأقل من 105 هو:

$$a.m. = \frac{105+100}{2} = 102.5$$

حيث أن :

a.m. - هو مركز الفئة (Arithmetic Mean)

جدول (4 - 1)

جدول التوزيع التكراري لدرجات الذكاء للمهندسين الخمسين

التكرار (f _i)	العلامات	الفئات (Sets)
1	/	- 90
2	//	- 95
8	/// ###	- 100
6	/ ###	- 105
6	/ ###	- 110
10	### ###	- 115
7	// ###	- 120
4	////	- 125
3	///	- 130
1	/	- 135
2	//	140 إلى أقل من 145
50		المجموع (Total)

وقد يكون من المناسب عرض البيانات في بعض الأحيان على شكل توزيع تكراري نسبي (Relative Frequency Distribution R.F.D.) أي إظهار تكرار كل فئة كنسبة من المجموع الكلي للتكرارات . والجدول (5-1) يبين

تبويب البيانات الخاصة بدرجات الذكاء للخمسين مهندس المختارين في العينة العشوائية السابقة في شكل جدول توزيع تكراري نسبي مكوناً من ست فئات طول كل فئة منها عشرة درجات .

جدول (5 - 1)

الفئات (Sets)	التكرار (f_i)	التكرار النسبي (p)	التكرار المئوي ($P\%$)
90 -	3	0.06	6 %
100 -	14	0.28	28 %
110 -	16	0.32	32 %
120 -	11	0.22	22 %
130 -	4	0.08	8 %
140 وأقل من 145	2	0.04	4 %
المجموع (Total)	50	1.00	100 %

نلاحظ من الجدول (4-1) والجدول (5-1) أنه تم استخدام فئات ذات أطوال متساوية ، يسمى التوزيع التكراري من هذا النوع بالتوزيع التكراري المنتظم أو ما يدعى (Uniform Frequency Distribution) .

وتجدر الإشارة هنا إلى أنه من الأفضل دائماً استخدام الفئات متساوية الطول ، وذلك تسهيلاً للعمليات الحسابية غير أنه في بعض الأحيان قد نضطر لاستخدام فئات غير متساوية الطول وهنا يسمى التوزيع التكراري بالتوزيع التكراري غير المنتظم (Non-Uniform Frequency Distribution) .

كما ويلاحظ من الجدول (5-1) أن بداية الفئة الأولى ونهاية الفئة الأخيرة محددين لذلك يسمى هذا النوع من الجداول المقفلة (Closed Tables) أما

الجداول التي تكون فيها نهاية الفئة الأخيرة غير محددة فتسمى بالجداول المفتوحة Open Tables.

1.4.6.1 جداول التوزيعات التكرارية المزدوجة (Double Frequency Distribution Tables)

لاحظنا في الجدول (1-3) كيفية تكوين الجداول التكرارية البسيطة وذلك بتبويب مجموعة من المفردات التي تخص ظاهرة واحدة ، أما إذا كان لدينا مجموعة من أزواج القيم لمتغيرين تربطهما علاقة معينة مثل مجموعة من الأزواج وأعمار زوجاتهم أو أوزان أشخاص ما وأطوالهم وغيرها ، فإننا نستطيع تبويب مثل هذه البيانات في جدول تكراري مزدوج أو ما يصطلح على تسميته (Double Frequency Table) ، وهو جدول ينقسم أفقياً ورأسياً إلى عدد من الصفوف والأعمدة بحيث يبين التقسيم الرأسي فئات إحدى الظاهرتين ويبين التقسيم الأفقي فئات الظاهرة الأخرى .

فمثلاً إذا أردنا تبويب البيانات الخاصة بكل من درجات اختبار الذكاء ودرجات اختبار المهارة للمهندسين الخمسين في العينة العشوائية السابقة في شكل جدول تكراري مزدوج ، فإننا نقسم مدى كل من المتغيرين إلى عدد مناسب من الفئات ثم نقوم بتفريغ البيانات ، بأن نضع لكل قيمتين متناظرتين علامة (/) في الخلية التي تقابل فئتيهما . فلو قسمنا المدى لدرجات الذكاء إلى ست فئات طول كل منها 10 درجات ، وقسمنا المدى لدرجات المهارة إلى خمسة فئات طول كل منها 10 درجات أيضاً ، بأن نضع لكل زوج من أزواج القيم التي لدينا والتي تخص أحد المهندسين علامة في الخلية التي ينتمي إليها هذا المهندس .

فالمهندس الأول حصل على 91 درجة في اختبار الذكاء و56 درجة في اختبار المهارة فهو إذن ينتمي إلى الخلية التي تقع في الصف الأول والعمود الثاني حيث أن درجة ذكاء هذا المهندس تقع في الفئة

(90 - أقل من 100) ودرجة مهارته تقع في الفئة (50 - أقل من 60) ونستمر في عملية التفريغ هذه مكونين عدداً من الحزم والعلامات كما تم ذكره سابقاً عند توضيح تكوين الجداول التكرارية البسيطة .

جدول (1 - 6)

تفريغ درجات اختبار الذكاء والمهارة للخمسين مهندس

درجات اختبار المهارة	درجات اختبار الذكاء	40 -	50 -	60 -	70 -	80 وأقل من 90
90 -	///	/				
100 -	///	///	///			
110 -		///	///	///		
120 -			///	///	///	
130 -					///	/
140 وأقل من 150						///

وبعد الانتهاء من التفريغ يتم استبدال العلامات بعددها فنحصل على الجدول التكراري المزدوج الذي يبين توزيع المهندسين حسب الدرجات التي حصلوا عليها في اختباري الذكاء والمهارة كما يبين الجدول (1-5) ، حيث يلاحظ في العمود الأول والأخير يعطيان التوزيع التكراري لدرجات الذكاء بينما يعطي الصف الأول والصف الأخير التوزيع التكراري لدرجات المهارة . لذلك يجب أن يكون مجموع كل من العمود الأخير والصف الأخير مساوياً لعدد المهندسين في العينة العشوائية المدروسة .

جدول (1-7)

جدول التوزيع التكراري لدرجات اختباري الذكاء والمهارة

فئات درجات اختبار الذكاء فئات اختبار درجات المهارة	40-	50 -	60 -	70 -	80 وأقل من 90	المجموع
90 -	2	1				3
100 -	6	8				14
110 -		4	10	2		16
120 -		3	4	4		11
130 -				3	1	4
140 وأقل من 150					2	2
المجموع	8	16	14	9	3	50

وفي كثير من الأحيان قد نحتاج لمعرفة عدد من المفردات التي تقل قيمتها عن حد معين أو تلك التي تساوي قيمتها أو تزيد عن حد معين ففي الجدول (1-3) ، قد نحتاج لمعرفة عدد المهندسين الذين حصلوا على درجات في اختبار الذكاء تقل عن 110 درجة ، وهؤلاء عددهم 17 مهندس لأنهم يمثلون مجموع تكرار الفئتين الأولى والثانية وهما (90 إلى أقل من 100) و (100 إلى أقل من 110) وهو $3 + 14 = 17$ مهندس ، بينما إذا أردنا معرفة عدد المهندسين الذين حصلوا على درجات تساوي أو تزيد عن 130 درجة فنجدهم 6 مهندسين لأنهم يمثلون مجموع تكراري الفئتين الأخيرتين وهما (130 وأقل من 140) والفئة (140 وأقل من 150) وهو $4 + 2 = 6$ مهندسين .

ولمعرفة هذا النوع من البيانات نكون ما يسمى بجدول التكرار المتجمع (Cumulative Frequency Table) ، جدول التكرار المتجمع النسبي (Relative Cumulative Frequency Table) ، وهناك نوعين من الجداول المتجمعة المستخدمة بكثرة هي جدول التكرار المتجمع الصاعد (Ascending Cumulative Frequency) ، وجدول التكرار المتجمع النازل (Descending Cumulative Frequency) .

إن الجدول (8-1) ، يمكننا من معرفة عدد المهندسين الذين حصلوا على درجات تقل عن الحدود العليا للفئات ونحصل عليه بتجميع بيانات التكرارات من الجدول (3-1) من جهة الفئات الصغيرة إلى الكبيرة .

جدول (8 - 1)

التكرار المتجمع الصاعد لدرجات الذكاء للخمسين مهندس

الحدود العليا للفئات	التكرار المتجمع الصاعد	التكرار المتجمع الصاعد النسبي
أقل من 90	0	0
أقل من 100	3	0.06
أقل من 110	17	0.34
أقل من 120	33	0.66
أقل من 130	44	0.88
أقل من 140	48	0.96
أقل من 150	50	1.00

أما الجدول (9-1) فيبين لنا عدد المهندسين الذين حصلوا على درجات تساوي أو تزيد عن الحدود الدنيا للفئات ونحصل عليه بتجميع بيانات

التكرارات من الجدول (3-1) ابتداءً من الفئات الكبيرة إلى الفئات الصغيرة أي يرمز للفئة المعنية بعدها الأعلى فأكثر وهكذا .

جدول (1 - 9)

التكرار المتجمع النازل لدرجات اختبار الذكاء للخمسين مهندس

الحدود الدنيا للفئات	التكرار المتجمع النازل	التكرار المتجمع النازل النسبي
90 فأكثر	50	1.00
100 فأكثر	47	0.94
110 فأكثر	33	0.66
120 فأكثر	17	0.34
130 فأكثر	6	0.11
140 فأكثر	3	0.04
150 فأكثر	0	0

7.1 التمثيل والعرض البياني للبيانات

(Graphical Representation of Data)

يفيد العرض البياني في إظهار البيانات العددية وتتبع المتغيرات فيها بطريقة تجلب الانتباه وتتسم بالبساطة والسهولة في تذكرها ، كما تفيد أيضا في توضيح العلاقات بين المتغيرات التي ندرسها .

وتختلف الرسائل التي نستخدمها تبعاً لنوع البيان الإحصائي والحقائق المطلوب إبرازها، ونذكر فيما يلي بعض الطرق التي تستخدم في العرض والتمثيل البياني :

1- الخط البياني (Graphical Line)

يستخدم الخط البياني لتمثيل العلاقة بين ظاهرتين أو متغيرين ، بحيث يبين كيفية تغير إحدى الظاهرتين مع الظاهرة الأخرى أو تبعاً لها . فإذا كان الزمن هو أحد المتغيرين ، يكون الغرض من الرسم هو معرفة مدى التغير الذي يحدث في الظاهرة التي ندرسها خلال فترة زمنية محددة . وفي هذه الحالة يسمى الخط البياني الذي نحصل عليه بالمنحنى التاريخي للظاهرة (Historical Curve) .

2- الأعمدة البيانية (Bar Charts)

وهي أعمدة أو مستطيلات رأسية قواعدما متساوية لأنها تعتمد على طول الفئة أو الفترة التي غالباً ما تكون متساوية ، أما ارتفاعاتها فتتناسب مع قيمة الظاهرة التي ندرسها أي التكرار (f_i) فإذا كان المحور الأفقي يمثل الزمن فإن التغير في ارتفاع المستطيلات يمثل التطور التاريخي للظاهرة ، وقد تستخدم الأعمدة البيانية للمقارنة بين أكثر من ظاهرة وذلك برسم أعمدة متلاصقة للظواهر المراد مقارنتها في السنوات المختلفة على أن يستخدم لكل ظاهرة لون مختلف أو ظل مختلف .

وفي بعض الأحيان قد يوضع فرق بين الأعمدة لتسهيل المقارنة وذلك إذا كانت البيانات التي لدينا هي بيانات إحصائية إجمالية مقسمة إلى مكوناتها ، مثل بيانات عدد السكان في السنوات المختلفة مقسمة إلى ذكور وإناث وهنا نرسم عموداً يمثل عدد السكان لكل سنة والجزء الأسفل منه يمثل الذكور والجزء الأعلى يمثل الإناث ، وإذا استخدمنا الأعمدة البيانية لدراسة ظاهرة معينة أو عدد من الظواهر بحيث لا تشتمل على عنصر زمني كأن نقسم الطلبة حسب المراحل الدراسية المختلفة أو العمال حسب الصناعات

المختلفة ، فيستحسن في هذه الحالة ترتيب الأعمدة حسب قيمها ترتيباً تصاعدياً أو ترتيباً تنازلياً لأن ذلك يبسط الشكل ويسهل من دراسته . ولاحظ أيضاً أنه إذا كان أحد الأعمدة أطول بكثير من الأعمدة الأخرى بحيث لا تتسع مساحة الرسم فإنه يمكن كسر العمود قبل نهايته مع توضيح القيمة التي تمثله في أعلى العمود ويمكن استخدام قضبان أفقية بدلاً من الأعمدة الرأسية وفي هذه الحالة يستبدل المحورين الأفقي والرأسي في جميع الحالات التي أشرنا إليها .

3- البيانات أو القطاعات الدائرية (Pie Charts)

تستخدم الدوائر عادةً للمقارنة بين المكونات المختلفة لظاهرة معينة ببعضها البعض وبين كل منها والمجموع . وبذلك تظهر الأهمية النسبية لهذه المكونات ويتم ذلك بتقسيم الدائرة إلى عدد من القطاعات تتلاقى في المركز بحيث تتناسب مساحة هذه القطاعات مع القيم المختلفة للمكونات الجزئية للظاهرة . وحيث أن الزاوية المركزية الرئيسية للدائرة هي 360° ، فإن القطاع الذي تكون زاويته 3.6° سوف يمثل 1% من مساحة الدائرة وبذلك يمكن تحديد أي قطاع بواسطة زاويته المركزية ، فمثلاً القطاع الذي تكون مساحته 30% من مساحة الدائرة تكون زاويته المركزية هي :

$$360^\circ \times \frac{30}{100} = 72^\circ$$

والقطاع الذي تمثل مساحته 40% من مساحة الدائرة ستكون زاويته المركزية 169.2° وهكذا .

ولتوضيح كيفية التمثيل بهذه الطريقة نقوم بدراسة المثال التالي .

مثال (1-1)

الجدول (10-1) يبين مساحات القارات المختلفة في العالم مقدره بملايين الكيلومترات المربعة والمطلوب تمثيل هذه البيانات بالآتي :

(a) القطاعات الدائرية .

(b) الأعمدة البيانية .

جدول (10-1)

أمريكا الجنوبية	أمريكا الشمالية	أوروبا	آسيا	أستراليا	أفريقيا
18	24	11	44	8	30

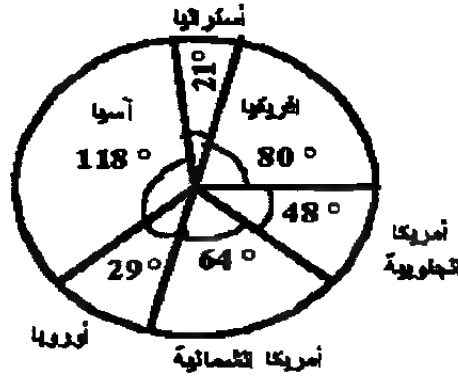
الحل :

نقوم أولاً بإيجاد الزوايا المقابلة والتي تمثل مساحة كل قطاع والموضحة في الجدول (11-1) ، بما أن الزاوية المركزية الرئيسية للدائرة هي 360° إذن :

جدول (11-1)

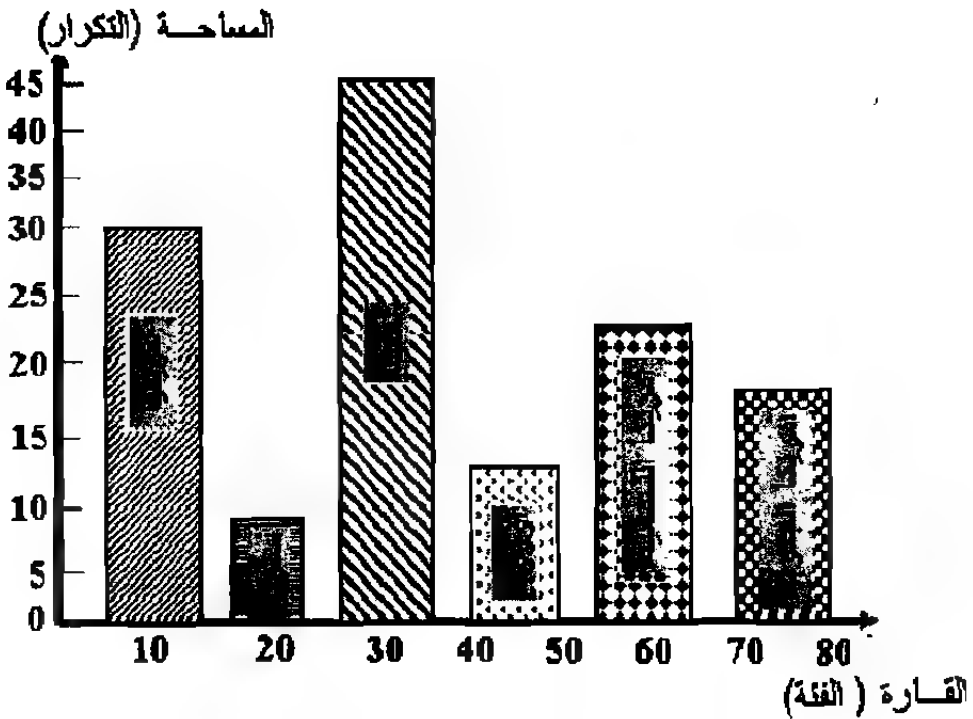
ت	الفئة (القارة)	المساحة (التكرار)	زاوية القطاع
1	أفريقيا	30	$80^\circ = 360^\circ \times \frac{30}{135}$
2	أستراليا	8	$21^\circ = 360^\circ \times \frac{8}{135}$
3	آسيا	44	$118^\circ = 360^\circ \times \frac{44}{135}$
4	أوروبا	11	$29^\circ = 360^\circ \times \frac{11}{135}$
5	أمريكا الشمالية	24	$64^\circ = 360^\circ \times \frac{24}{135}$
6	أمريكا الجنوبية	18	$48^\circ = 360^\circ \times \frac{18}{135}$
	المجموع	135	360°

بعد ذلك نقوم برسم دائرة بنصف قطر مناسب ونقسم زاويتها المركزية حسب نسبة القطاعات المختلفة والزوايا المقابلة لها فنحصل على الشكل (1-1) :



الشكل (1-1)

أما بالنسبة لتمثيل البيانات بطريقة الأعمدة البيانية فتكون كما يوضح الشكل (2-1) ، حيث تمثل القارات على المحور الأفقي بفترات متساوية ويمثل التكرار بمستطيل .



الشكل (2-1)

8.1 التمثيل البياني للتوزيعات التكرارية

(Graphical Representation of Frequency Distributions)

يمكن تمثيل التوزيعات التكرارية بيانياً بعدة طرق مختلفة أهمها المدرج التكراري (Histogram) ، والمضلع التكراري (Frequency Polygon) ، وكذلك المنحني التكراري (Frequency Curve) ، وسوف نقوم بعرض طرق التمثيل البياني لهذه التوزيعات بالتفصيل :

1- المدرج التكراري (The Histogram)

في هذا التمثيل يتم إنشاء محورين وهما المحور الأفقي والمحور الرأسي ، وتمثل الفئات أو الفترات على المحور الأفقي والتكرارات على المحور الرأسي ، ويراعى عند التمثيل البياني اختيار مقياس رسم مناسب لكل محور ، بحيث يكفي المحور الأفقي لتمثيل جميع الفئات ويسمح المحور الرأسي بتمثيل جميع التكرارات ثم نكوّن المدرج التكراري وذلك برسم عدد من المستطيلات المتلاصقة بحيث يكون مساحتها متناسبة مع التكرارات .

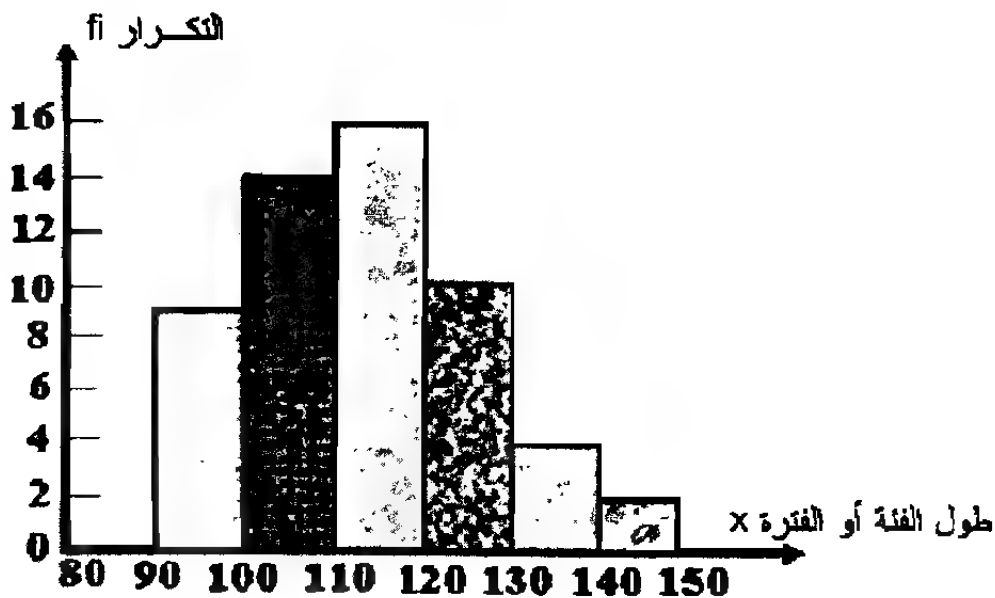
إذا كانت أطوال الفئات متساوية فإننا ننشئ عدداً من المستطيلات قاعدة كل مستطيل تمثل طول الفئة ، وارتفاعه هو تكرار تلك الفئة وهنا سوف نجد بأن مساحة المستطيلات تتناسب مع التكرارات لأن أطوال قواعد المستطيلات جميعها متساوية ولأن مساحة كل مستطيل تساوي طول القاعدة في الارتفاع .

ويلاحظ هنا أن المحور الرأسي لا بد وأن يبدأ من الصفر حتى يمكننا مقارنة التكرارات المختلفة ببعضها البعض ، أما المحور الأفقي فليس من الضروري أن يبدأ من الصفر حيث لو فعلنا ذلك لاحتجنا إلى مسافة كبيرة جداً لتمثيل الفئات بينما يكون المحور الأفقي من الصفر حتى بداية الفئة الأولى غير مستعمل .

ولو عدنا إلى الجدول (1-4) والذي يمثل التوزيع التكراري لدرجات الذكاء للخمسين مهندس لوجدنا بأنه من الممكن تمثيل تلك البيانات على شكل مدرج تكراري ، حيث يتبين أن أطوال الفئات هي متساوية لأن طول كل فئة يساوي 10 درجات ، لذا فأننا نقسم المحور الأفقي إلى ستة أقسام متساوية تمثل فئات الدرجات .

أما المحور الرأسي فيتم تقسيمه إلى أقسام متساوية بحيث يسمح بتمثيل أكبر التكرارات وهو 16 في مثالنا هذا ثم نرسم ستة مستطيلات طول قاعدة كل منها 10 درجات وارتفاع كل منها يساوي تكرار الفئة المقام فوقها هذا المستطيل ، وبذلك نحصل على المدرج التكراري الموضح في الشكل (1-3) والذي يبين أن مساحة المستطيلات تتناسب مع التكرارات .

وفي حالة الجداول غير المنتظمة أي ذات الفئات غير المتساوية الأطوال فإنه يلزم تعديل التكرارات قبل رسم المدرج التكراري ، وذلك بقسمة تكرار كل فئة على طول هذه الفئة وبذلك نحصل على التكرار المعدل ثم نكون المدرج التكراري بإنشاء عدد من المستطيلات قواعدها هي أطوال الفئات غير المتساوية ارتفاعاتها هي التكرارات المعدلة وفي هذه نجد أن مساحة كل مستطيل تساوي تماماً تكرار الفئة التي يمثلها .



الشكل (1 - 3)

تمثيل المدرج التكراري لدرجات الذكاء للخمسين مهندس

وكمثال على ذلك دعنا نرسم المدرج التكراري لتوزيع المصانع المبين في الجدول (1- 12) .

جدول (1 - 12)

توزيع 200 مصنع حسب عدد العمال الذين يستخدمهم كل مصنع

فئات العمال x	- 1	- 3	- 5	- 10	- 20	50 وأقل من 100
عدد المصانع f	8	22	80	48	27	15

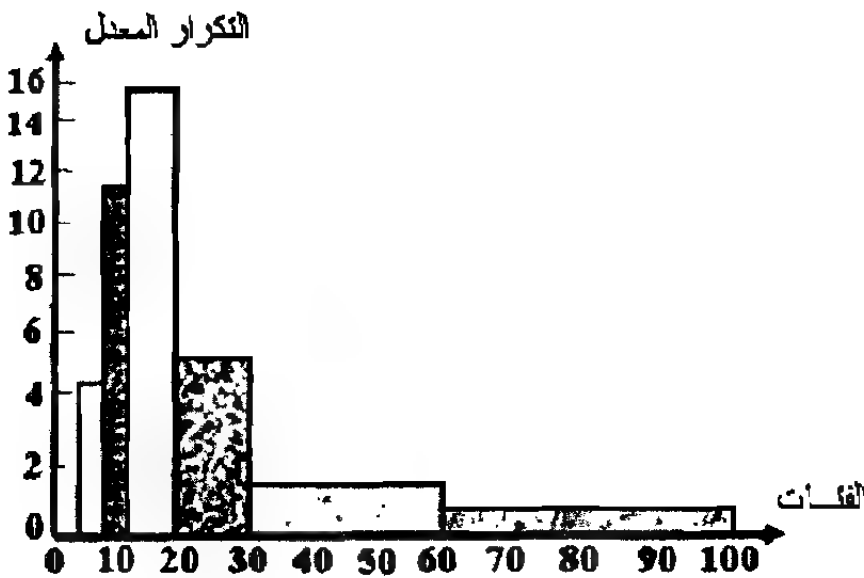
ويلاحظ في هذا التوزيع عدم تساوي أطوال الفئات ولذلك يلزم تعديل التكرارات قبل رسم المدرج التكراري كما موضح في الجدول (1- 13)

وبتمثيل الفئات على المحور الأفقي والتكرار المعدل على المحور الرأسي نحصل على المدرج التكراري الموضح في الشكل (1 - 4).

جدول (13-1)

ت	الفئات	التكرار	طول الفئة	التكرار المعدل
1	1 -	8	2	4
2	3 -	22	2	11
3	5 -	80	5	16
4	10 -	48	10	4.8
5	20 -	27	30	0.9
6	50 وأقل من 100	15	50	0.3
المجموع Total		200		—

وعند النظر للشكل (1-4) سنجد أن مساحة كل من المستطيلات تساوي حاصل ضرب القاعدة أي طول الفئة في الارتفاع أي التكرار المعدل والتي تساوي التكرار الأصلي .

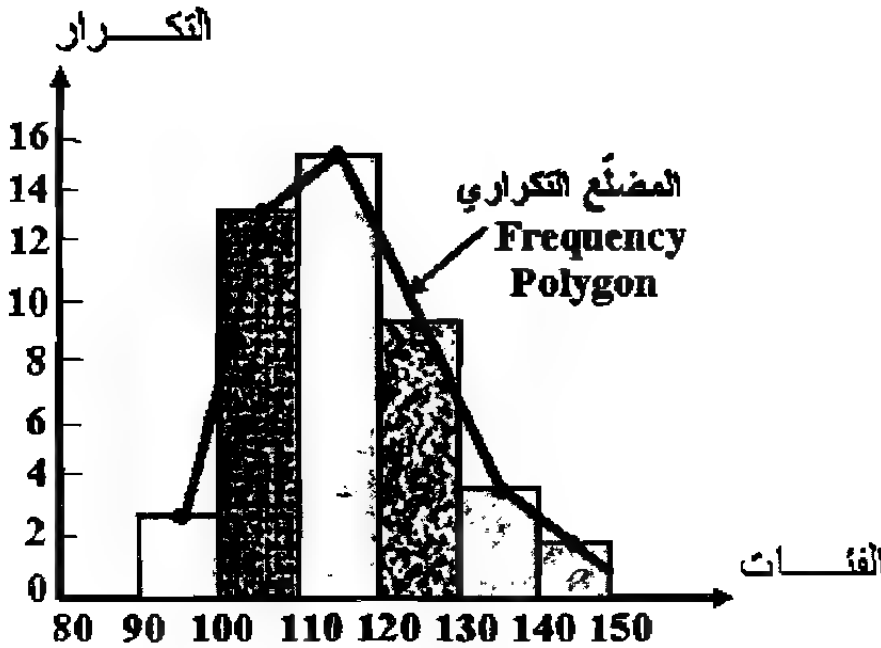


الشكل (1 - 4)

المدرج التكراري لفئات عمال المصانع غير المتساوية

2-المضلع التكراري (The Frequency Polygon)

سبق أن أوضحنا عند تكوين الجداول التكرارية انه يمكن اعتبار مركز الفئة ممثلاً لجميع القيم الواقعة في هذه الفئة ، فلو مثلنا مراكز الفئات على المحور الأفقي والتكرارات على المحور الرأسي فإنه يمكن تمثيل بيانات الجدول التكراري بعدد من النقاط تقع كل منها أمام مركز الفئة على بعد رأسي تكرار هذه الفئة ، وبتوصيل هذه النقاط بخطوط مستقيمة نحصل على المضلع التكراري كما موضح في الشكل (1-5) ، الذي يبين المضلع التكراري لتوزيع درجات الذكاء للخمسين مهندس والمدرجة في الجدول (1 - 4) .

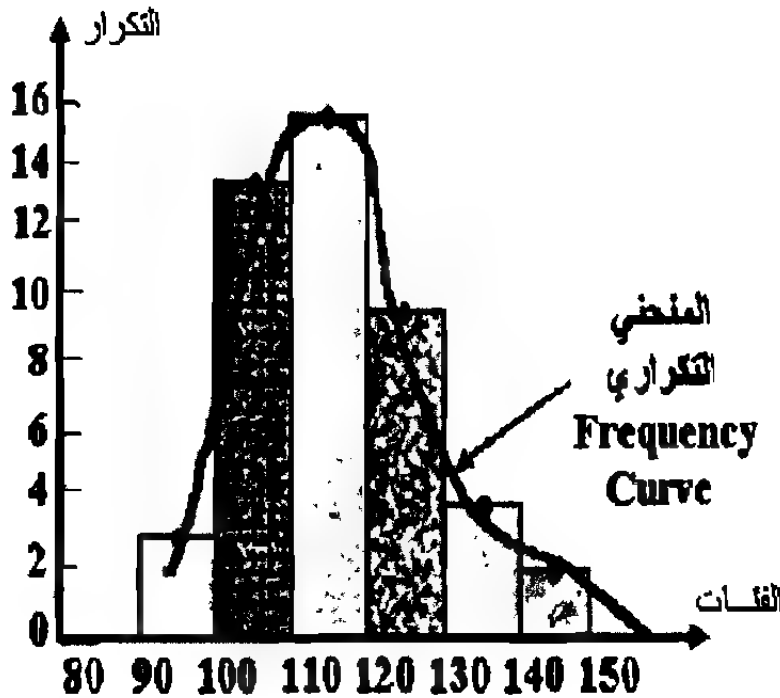


الشكل (1 - 5)

المضلع التكراري لتوزيع درجات الذكاء للخمسين مهندس

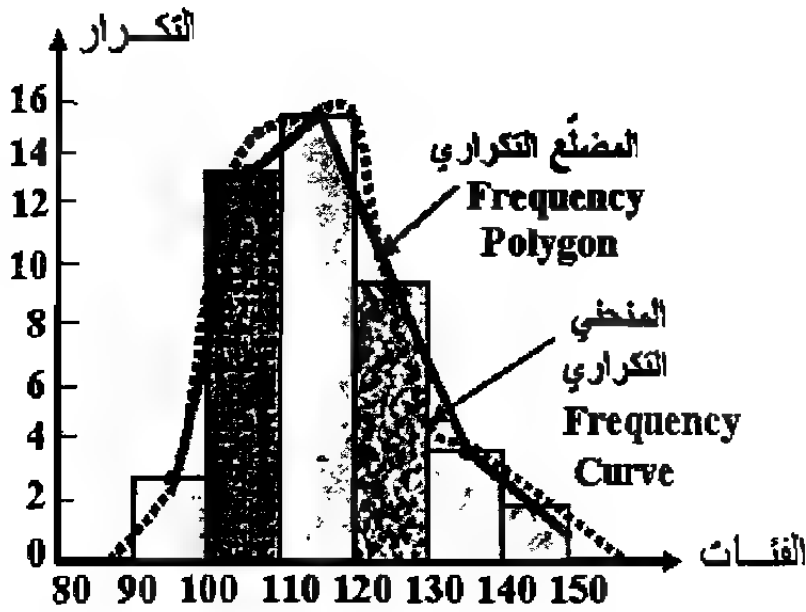
3- المنحنى التكراري (The Frequency Curve)

بتمثيل مراكز الفئات على المحور الأفقي والتكرارات على المحور الرأسي نعين عدداً من النقاط ولتكن مراكز الفئات ، ونحاول أن نرسم منحنى بينها بدلاً من الخط المنكسر في حالة المضلع التكراري على أن يمر بمعظم النقاط ويتوسط بقيتها خير توسط ، وهذا المنحنى يعرف بالمنحنى التكراري . والشكل (6-1) يبين منحنى التوزيع التكراري لدرجات الذكاء للخمسين مهندساً أما الشكل (7-1) فيبين رسم كل من المضلع التكراري والمنحنى التكراري في نفس الإحداثيات حتى يمكن المقارنة بينهما حيث يتضح أن المنحنى التكراري لا يمر بجميع نقاط المضلع .



الشكل (6 - 1)

المنحنى التكراري لدرجات الذكاء للخمسين مهندساً



الشكل (1 - 7)

مقارنة بين المضلع التكراري والمنحنى التكراري
لتوزيع درجات الذكاء للخمسين مهندس

وبلاحظ أنه كلما زاد عدد المفردات وقصرت أطوال الفئات فإن الخط الذي يمر برؤوس المضلع يكون أقل انكساراً وأكثر تمهيداً وسيقترب المدرج التكراري من المنحنى التكراري .

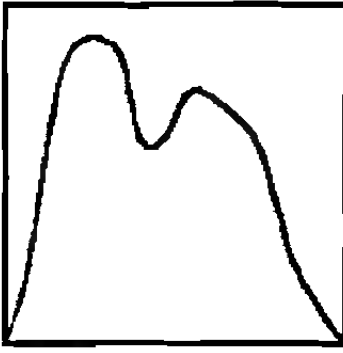
4- أشكال المنحنيات التكرارية (Frequency Curves Types)

تأخذ المنحنيات التكرارية أشكالاً مختلفة كما يتضح من الشكل (1-8) وأكثر هذه المنحنيات شيوعاً هو المنحنى المعتدل (Normal Curve) ، وهو منحنى متماثل (Symmetric) يشبه الناقوس وله نهاية عظمى في منتصفه تماماً ويتمثل حول محور رأسي يمر بهذه النهاية العظمى بحيث يقسم المساحة تحت المنحنى إلى قسمين متطابقين ولهذا المنحنى أهمية كبيرة في الدراسات الإحصائية ولذلك سندرسه بالتفصيل فيما بعد .

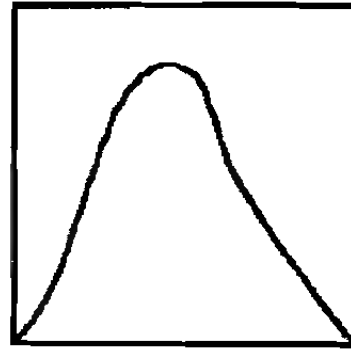
والمنحنى المعتدل هو الشكل الذي نتوقع الحصول عليه من دراسة كثير من الظواهر التي تتغير تبعاً لأسباب طبيعية مثل الطول والوزن وغيرها ، وفي بعض الحالات قد نحصل على أشكال قريبة من المنحنى المعتدل ولكنها ليست متماثلة تماماً فيكون للمنحنى قمة واحدة ولكن أحد طرفيه أطول من الطرف الآخر ويسمى ذلك بالالتواء (Skewness) ويكون الالتواء موجباً إذا كان الطرف الأيمن أو الذيل الأيمن للمنحنى هو الأطول (Positive Skewness) ، أما عندما يكون الطرف الأيسر أو الذيل الأيسر للمنحنى هو الأطول فيسمى ذلك الالتواء سالباً (Negative Skewness) .

وفي حالات أخرى قد نجد أن المنحنى التكراري يكون ذا "شعبة واحدة" أو ما يدعى (Reverse J - Shaped Curve) ، ويحدث ذلك عند وجود أكبر التكرارات ومن أمثلة ذلك توزيع ملكية الأراضي حسب فئات المساحة المملوكة حيث أن عدد الملكيات الصغيرة كبير جداً بينما يقل عدد الملاك كلما كبرت المساحة المملوكة . وبذلك يكون للمنحنى فرع واحد أي شعبة واحدة كما في الشكل (8-1) .

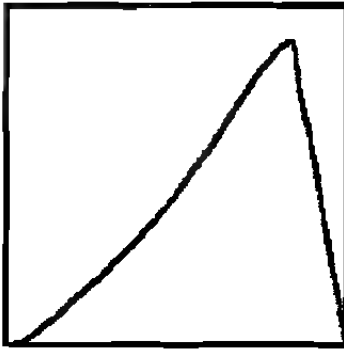
أما إذا وجدت أكبر التكرارات عند طرف المنحنى فإن المنحنى يكون ذو شعبتين (U - Shaped Curve) ويشابه في شكله حرف U ، ومن أمثلة ذلك توزيع الوفيات حسب فئات العمر فنجد أن عدد الوفيات يكون كبيراً في سنوات العمر الأولى ثم يقل في فترة الشباب ثم يرتفع مرة أخرى في فترة الشيخوخة وبذلك نحصل على منحنى ذو شعبتين .



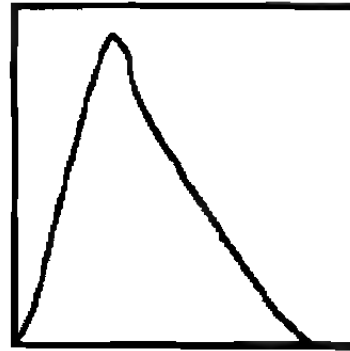
منحنى ذو قيمتين
Bimodal Curve



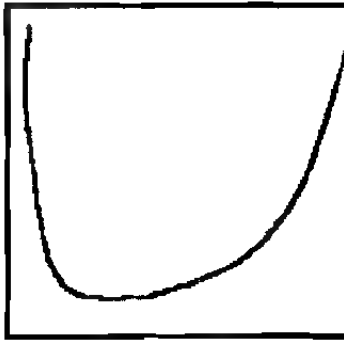
منحنى تكراري متماثل
Symmetric Frequency Curve



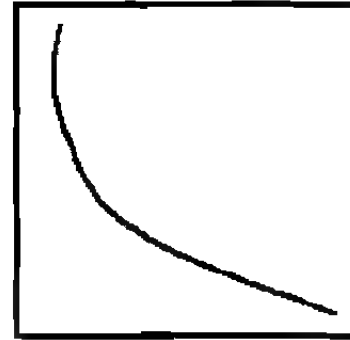
منحنى مائل الالىف
Negative Skewness Curve



منحنى موجب الالىف
Positive Skewness Curve



منحنى ذو شعبتين
U-Shaped Curve



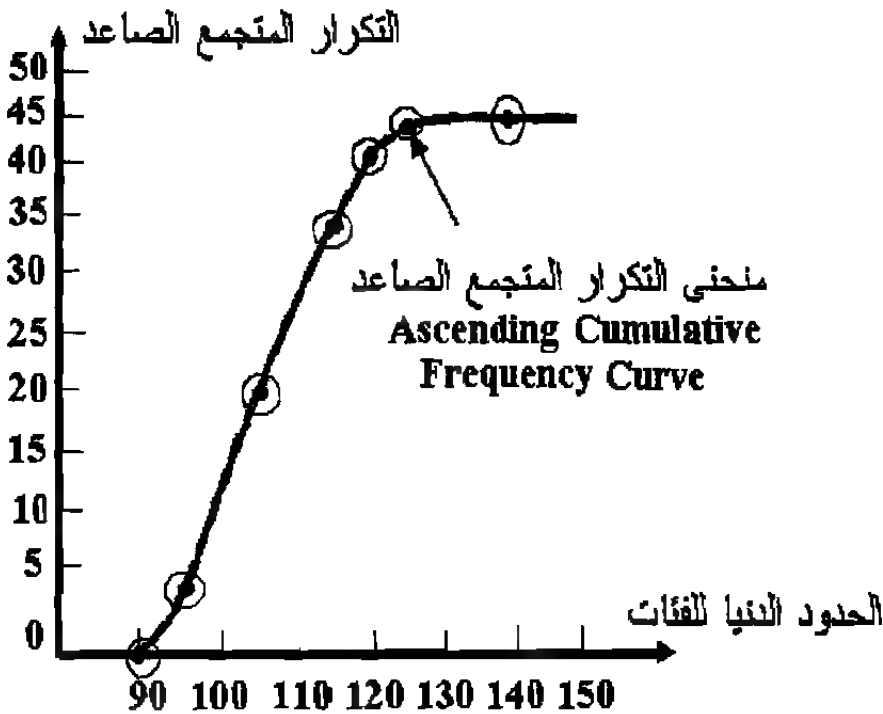
منحنى ذو شعبة واحدة
Reverse J-Shaped Curve

الشكل (1 - 8)

أشكال المنحنيات التكرارية

5- المنحنى التكراري المتجمع (Cumulative Frequency Curve)

عند دراستنا لتبويب البيانات أوضحنا كيفية تكوين جداول التكرار المتجمع الصاعد والنازل . ولتمثيل هذه الجداول بيانياً ، نمثل حدود الفئات على المحور الأفقي والتكرارات المتجمعة على المحور الرأسي فنحصل على كل من المنحنى المتجمع الصاعد والمنحنى المتجمع النازل وفي الشكل (9-1) تم تمثيل منحنى التكرار المتجمع الصاعد لتوزيعات درجات الذكاء للخمسين مهندس وفي الشكل (10-1) تم تمثيل منحنى التكرار المتجمع النازل لدرجات الذكاء لنفس الخمسين مهندس .



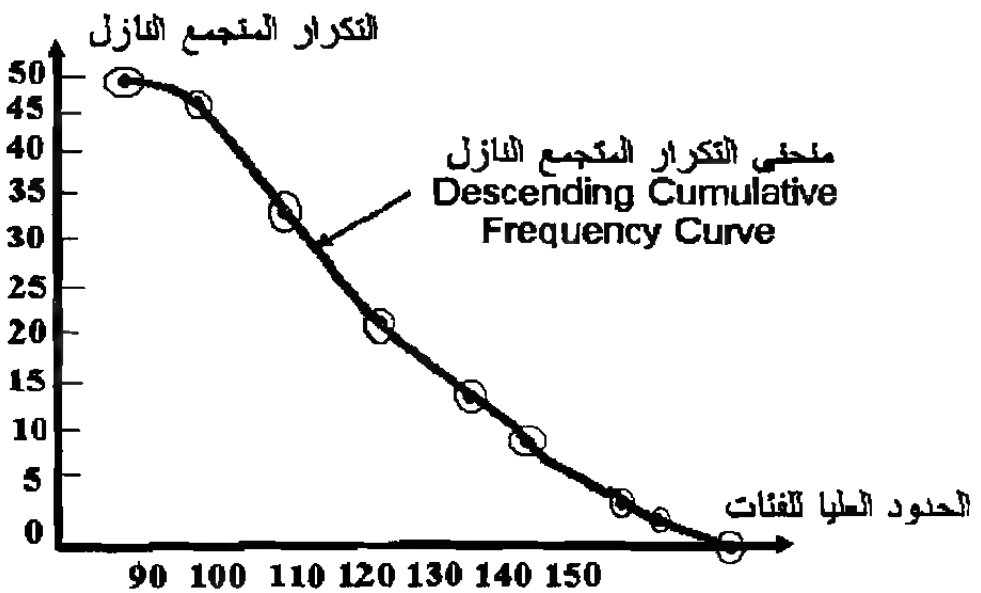
الشكل (9 - 1)

منحنى التوزيع المتجمع الصاعد لدرجات الذكاء للخمسين مهندس

ويتضح من الشكل أعلاه أن عدد المهندسين الذين حصلوا على درجات أقل من 110 هو 17 مهندس أي بنسبة 34% من مجموع المهندسين وأن عدد المهندسين الذين حصلوا على درجات أقل من 130 هو 33 مهندس أي بنسبة 66% من مجموع المهندسين وهكذا . وبذلك يمكننا استخدام الشكل في معرفة عدد المهندسين الذين حصلوا على درجات أقل من حد معين وبالتالي معرفة النسبة المئوية لهؤلاء المهندسين ، فلو أردنا معرفة النسبة المئوية للمهندسين الذين حصلوا على درجات أقل من 135 درجة نحسب أولاً عددهم وذلك بإقامة عمود على المحور الأفقي عند القيمة 135 حتى يلاقي المنحني المتجمع الصاعد في نقطة ثم نرسم منها خطاً أفقياً موازياً للمحور الأفقي حتى يقابل المحور الرأسي في نقطة تحدد لنا عدد هؤلاء المهندسين فنجد أنه مثلاً 46 وهؤلاء تكون نسبتهم :

$$100 \times \frac{46}{50} = 92 \%$$

مساوية إلى 92 % من مجموع المهندسين .



الشكل (1 - 9)

منحني التوزيع المتجمع النازل لدرجات الذكاء للخمسين مهندس

9.1 تمارين (Exercises)

س1: الجدول (14-1) يبين عدد الوحدات التي تم سحبها من مخزون أحد المصانع خلال فترة 40 يوماً :

جدول (14-1)

99	83	97	87	82	88	81	91	80	83
98	93	78	87	90	84	92	73	85	75
82	89	101	82	83	88	86	93	80	86
94	103	81	76	92	84	89	80	95	85

المطلوب :

- (a) تفرغ هذه البيانات في جدول تكراري مناسب .
(b) رسم المدرج التكراري والمضلع التكراري للبيانات .

س2: الجدول (15-1) يبين عينة عينة عشوائية مكونة من 25 عاملاً أخذت من عمال أحد المصانع فوجد أن متوسط الوحدات التي أنتجها هؤلاء العمال في الأسبوع كانت كما يلي :

جدول (15-1)

132	143	171	151	134
160	123	144	145	141
135	156	114	167	140
155	138	147	130	168
138	150	159	130	150

المطلوب :

- (a) إدراج البيانات أعلاه في جدول تكراري مناسب .
(b) رسم المدرج التكراري والمضلع التكراري للبيانات .
(c) مثل منحني التكرار المتجمع الصاعد للبيانات ومنه أستنتج نسبة العمال الذين نقل إبنانية كل منهم عن 135 وحده أسبوعياً .

س3: الجدول (16-1) يبين عدد الأسر لأقرب مائة والمقيمة في البلديات التالية حسب النتائج الأولية لتعداد 1973 للسكان في ليبيا .

جدول (16-1)

البلدية	أجدابيا	جالدو	بني وليد	سرت	الكفرة	المجموع
عدد الأسر	91	15	12	39	32	179

المطلوب تمثيل البيانات أعلاه بطريقة القطاعات الدائرية .

س4: الأعداد المبينة في الجدول (17-1) تمثل درجات 75 طالباً في امتحان لمادة مقاومة المواد في الفصل الثالث لإحدى الكليات الهندسية .

جدول (17-1)

68	84	75	82	68	90	62	88	76	93	73	79	85	73	60
83	71	59	85	75	61	65	75	87	74	62	91	78	63	72
66	75	82	75	84	77	69	74	68	60	96	78	79	63	75
90	60	79	83	71	79	62	69	97	78	85	76	65	71	75
80	53	73	57	68	78	66	62	65	74	76	67	73	81	82

المطلوب :

- (a) تمثيل البيانات في جدول تكراري مكون من 8 فئات .
 (b) تمثيل المدرج التكراري والمنحني التكراري لهذه البيانات .
 (c) رسم منحني التكرار للمتجمع الصاعد ومنحني التكرار للمتجمع النازل لهذه البيانات .

س5: البيانات في الجدول (18-1) تبين كمية الكهرباء المستهلكة يومياً بآلاف الكيلووات في إحدى المدن خلال فترة زمنية مقدارها 100 يوم .

جدول (18-1)

25	32	28	26	30	21	36	27	32	32	33	23	31
29	33	31	29	33	28	31	27	29	29	33	27	28
32	23	30	25	29	32	21	24	30	31	28	32	31
34	31	26	30	30	34	26	26	29	32	22	24	31
28	32	28	37	30	35	27	33	28	31	28	32	30
31	30	33	30	28	29	26	29	32	27	27	34	29
31	33	39	30	33	28	33	32	25	31	31	30	31
28	30	30	33	30	31	30	32	28				

المطلوب :

- (a) أكتب جدول التوزيع التكراري للبيانات ، اختر عدد الفئات 7 .
 (b) أنشئ المدرج التكراري والمضلع التكراري للبيانات .
 (c) أنشئ المدرج التكراري النسبي والمضلع التكراري النسبي للبيانات .
 (d) أكتب جدول التكرار للمتجمع الصاعد ثم جدول التكرار للمتجمع النازل لهذه البيانات .
 (e) مثل منحني التكرار للمتجمع الصاعد والنازل للبيانات .

س6: أراد عالم نفسي أن يجري تجربة على القردة ليتعرف من خلالها على إمكانية تعلم القردة أن الضوء الأحمر يعني قف والضوء الأخضر يعني تقدم ، فصار يكافئ القرد بموزة إذا استجاب الاستجابة الصحيحة للإشارة ويعاقبه بصدمة كهربائية إذا هو أخطأ ، أجرى التجربة على 100 قرد وسجل عدد المحاولات التي بذلها كل قرد حتى تعلم الدرس ، فحصل على النتائج المبينة في الجدول (19-1) :

جدول (19-1)

1	6	17	11	4	18	24	34	2	26
4	2	7	3	12	5	10	1	15	3
16	5	3	8	13	9	20	1	14	2
6	2	8	4	14	9	5	1	2	10
1	7	7	6	10	5	5	9	3	15
6	3	8	8	13	1	6	13	19	10
2	7	4	9	20	2	7	12	9	18
1	6	3	5	11	8	3	19	8	10
21	2	5	4	30	7	4	13	9	23
1	4	3	6	10	5	16	22	28	32

المطلوب :

- أكتب جدول التوزيع التكراري للبيانات أعلاه (أختر عدد الفئات 7) .
- أنشئ المدرج التكراري والمضلع التكراري للبيانات .
- أنشئ المدرج التكراري النسبي والمضلع التكراري النسبي للبيانات .
- أكتب جدول التكرار المتجمع الصاعد ثم جدول التكرار المتجمع النازل لهذه البيانات .
- مثل منحني التكرار المتجمع الصاعد والنازل للبيانات .

س7: جدول (20-1) يبين التوزيع التكراري لعدد الهكتارات التي يزرعها 150 مزارعاً في إحدى المناطق الزراعية :

جدول (20-1)

98 - 90	- 86	- 80	- 74	- 68	- 62	- 56	- 50	الفترة (الهكتار)
6	12	20	44	30	24	12	2	المزارعين (التكرار)

المطلوب :

- (a) مثل المدرج التكراري والمضلع التكراري والمنحني التكراري للبيانات .
 (b) أنشئ " جدول التكرار المتجمع الصاعد للبيانات ومثله بالرسم .
 (c) أنشئ " جدول التكرار المتجمع النازل للبيانات ومثله بالرسم .

س8: الجدول (21-1) يبين العمر بالشهر الذي تعلم فيه أربعون طفلاً ركوب الدراجة ذات العجلات الثلاثة :

جدول (21-1)

21	23	31	35	18	19	22	25	27	16
26	29	28	25	25	21	28	32	25	21
23	26	24	20	23	24	26	20	28	20
32	24	25	28	29	31	26	32	29	24

المطلوب:

- (a) نظم البيانات السابقة في جدول توزيع تكراري مكون من 5 فترات .

- (b) مثل التوزيع التكراري في (a) بمدرج تكراري ومضلع ومنحني تكراري .
(c) على نفس ورقة الرسم مثل التكرار المتجمع الصاعد والنازل للبيانات .

س9: الجدول (22-1) يبين الأجور الأسبوعية بالدينارين للعاملين بأحد شركات البترول :

جدول (22-1)

120 - 110	- 100	- 90	- 80	- 70	- 60	- 50	فئات الأجور (بالدينارين)
2	5	10	14	16	10	8	عدد العاملين (التكرار)

المطلوب :

(a) أرسم منحني التكرار المتجمع الصاعد للبيانات ومنه أستنتج نسبة العاملين الذين يحصلون على أجر أقل من 85 دينار أسبوعياً وعدد العاملين الذين يحصلون على أجر يتراوح بين 63 وأقل من 75 دينار أسبوعياً .

(b) أرسم منحني التكرار المتجمع النازل للبيانات ومنه أستنتج نسبة العاملين الذين يحصلون على أجر يساوي أو يزيد عن 98 دينار أسبوعياً .

س10: الجدول (23-1) يبين المساحة المزروعة قمحاً (\bar{x}) بآلاف الهكتارات وكمية الإنتاج من القمح (\bar{y}) بآلاف القنطارات والتي جمعت من 10 مناطق .

جدول (23-1)

المنطقة	1	2	3	4	5	6	7	8	9	10
\bar{x}	146	138	126	130	142	128	148	148	138	132
\bar{y}	999	993	996	987	993	1005	981	1008	1002	993

المطلوب :

تمثيل البيانات أعلاه في جدول تكراري مزدوج مناسب .

الباب الثاني

مقاييس النزعة المركزية

(Measures of Central Tendency)

- 1.2 مقدمة .
- 2.2 المقاييس الإحصائية للبيانات .
- 3.2 مقاييس النزعة المركزية .
- 4.2 المتوسط الحسابي .
- 5.2 طرق حساب المتوسط الحسابي .
- 6.2 حساب الوسط الحسابي باستخدام الوسط الفرضي .
- 7.2 حساب الوسط الحسابي بطريقة الانحرافات المختصرة .
- 8.2 خواص المتوسط الحسابي .
- 1.8.2 المتوسط الحسابي المرجح .
- 9.2 الوسيط .
- 1.9.2 خواص الوسيط .
- 2.9.2 طرق حساب الوسيط .
- 10.2 الربيع الأتني والربيع الأعلى .
- 11.2 إيجاد الوسيط والربيعين بيانياً .
- 12.2 خواص الوسيط .
- 13.2 المنوال .
- 1.13.2 الخواص العامة للمنوال .
- 2.13.2 طرق حساب المنوال .
- 14.2 الوسط الهندسي والوسط التوافقي والعشيرات والمثنيات .
- 1.14.2 الوسط الهندسي .
- 2.14.2 الوسط التوافقي .
- 3.14.2 العشيرات .
- 4.14.2 المثنيات .
- 15.2 العلاقة بين المتوسطات .
- 16.2 تمارين .

1.2 مقدمة (Introduction)

لاحظنا في الباب السابق أن الطرق البيانية لها أهميتها الخاصة عند تقديم المعلومات الإحصائية ، فهي تعطي وصفاً عاماً وسريعاً للبيانات الإحصائية ، ولكن فوائدها الاستقرارية قليلة . فالمضلع التكراري لعينة عشوائية مأخوذة من مجتمع ما يعطينا تصوراً عن شكل المضلع التكراري للمجتمع نفسه ، واستقراؤنا يقف عند الفرض بأنه يوجد تشابه ما بين مضلعين ولكن المشكلة التي نقف عندها هنا هي كيفية قياس مدى الاختلاف بينهما أو بصورة إيجابية قياس درجة التشابه بينهما . وهكذا يفضل وجود مقاييس وصفية أخرى يمكن استخدامها للتنبؤ بشكل توزيع التكرار الخاص بالمجتمع المدروس .

2.2 المقاييس الإحصائية للبيانات

(Statistical Measures of Data)

في كثير من التوزيعات التكرارية نجد أن عدداً كبيراً من المفردات يميل إلى التجمع حول قيمة متوسطة معينة والتي تمثل مركز التوزيع ويصطنح على تسمية هذه الظاهرة بالنزعة المركزية (Central Tendency) ، أو نزعة المفردات المختلفة إلى التجمع حول مركز التوزيع ، ويتضح من ذلك أن لكل مجموعة من البيانات قيمة متوسطة معينة خاصة بها تميزها عن مجموعات البيانات الأخرى والتي يمكن استخدامها لوصف الفئة أو المجموعة حيث أنها تحدد مركز أو متوسط الفئة .

وعلى سبيل المثال ، لو أخذنا عينة عشوائية من مجموع الطلاب الناجحين في شهادة الثانوية العامة الفرع العلمي للعام الدراسي (1980-1981) في مدينة عمان وسجلنا مجموع درجاتهم باستثناء مادة التربية الإسلامية فكانت على النحو المبين في الجدول (1-2) :

جدول (1-2)

117	109	210	125	105	190	143
173	130	125	221	145	105	137
103	111	180	105	93	112	106
101	117	172	102	98	107	94
102	98	104	153	103	99	112

من الدراسة سريعة لبيانات هذه العينة نلاحظ أنها تدل على ما يلي :

- (1) أكبر مجموع في هذه البيانات هو 221 .
- (2) المتوسط الحسابي لتلك البيانات هو 125.3 تقريباً .

إن العدد 221 يمثل مجموع درجات أحسن طالب في هذه العينة لكن المتوسط الحسابي 125.3 هو خاصية وصفية لهذه العينة . إن كلاً من هذين الرقمين سندعوه إحصائية في هذه العينة . وعلى هذا الأساس فإن أية قيمة عددية تصف عينة تدعى بالقيمة الإحصائية (Statistical Value) ، بينما أية قيمة عددية تصف مجتمعاً تدعى بالوسيط الإحصائي (Statistical Parameter) .

وهنا يجدر بنا الملاحظة بأننا يمكننا اختيار عدة عينات عشوائية من مجتمع ما وبالتالي نتوقع تغير الإحداثيات من عينة إلى أخرى . ففي مثالنا السابق إذا أخذنا عينة عشوائية أخرى من الطلاب الناجحين في مدينة عمان فسنجد أن المتوسط الحسابي للعينة الجديدة يختلف عن المتوسط الحسابي للعينة الأولى ، قد يختلف أحياناً ولكن هذه ليست بقاعدة كما أن الحد الأعلى للعينة أي مجموع درجات أحسن طالب فيها سيتغير أيضاً .

وستستخدم الإحصائيات التي نحصل عليها من عينة عشوائية مأخوذة من مجتمع ما لتقدير وسطاء هذا المجتمع . أي سنأخذ هذه الإحصائيات كقيم تقريبية لوسطاء المجتمع . حيث أن حجم المجتمع سيكون في أغلب الأحيان كبيراً . ولكن هل إن هذه الإحصائيات التي نحصل عليها لعينة ما تعتبر تعبيراً دقيقاً عن وسطاء المجتمع المأخوذة منه ، وكم هي درجة هذه الدقة ، وهل نكتفي بإحصائيات لعينة أم نلجأ لاختيار عدة عينات من المجتمع ، في الحقيقة سنلجأ لاختيار عدة عينات وسندرس توزيع الإحداثيات في هذه العينات المتعددة ونحدد درجة الدقة فيها .

3.2 مقاييس النزعة المركزية (Measures of Central Tendency)

ذكرنا في البند السابق بأنه عند دراسة مجموعة بيانات وصفية يكون من المفيد أن نعين مقاييس عددية إحصائيات أو وسطاء ، تصف الميزات الهامة لهذه البيانات وفي طليعة هذه البيانات العددية الهامة تأتي مقاييس النزعة المركزية أي القياسات المعبرة عن مواضع تتركز التوزيع (Measures of Location) .

وفي الكثير من التوزيعات التكرارية التي تصادفنا نجد بأنها تتسم بخواص عامة . فمثلاً القيم الصغرى يكون تكرارها قليلاً وبزيادة القيمة يكبر التكرار حتى يصل إلى نهاية عظمى ثم يعود التكرار إلى التناقص بكبر القيم الواردة بعد ذلك . كذلك فإننا نجد في معظم التوزيعات التكرارية أن عدداً كبيراً من الحالات تتجمع حول قيم في المدى الموزع فيه التكرار الكلي . أي أن عدداً كبيراً من الحالات تنجذب نحو قيمة معينة ، أي أن هذه القيمة تمثل مركز جذب لقيم الحالات الأخرى في التوزيع وتسمى مثل تلك الظاهرة بالنزعة المركزية .

فمثلاً ، إذا أجرينا لتلاميذ مدرسة ما اختباراً في الذكاء سوف نجد بأن معظمهم هم ذو ذكاء متوسط ، وأن قلة منهم تميل نحو العبقرية وقلة أخرى تميل نحو الغباء ، وإذا قمنا بقياس أطوال التلاميذ لأحد الفصول الثانوية في المدرسة فنجد بأن معظمهم يقتربون من طول معين وقلة منهم أكبر من هذا الطول وقلة أخرى أصغر منه ، وهكذا في الكثير من الظواهر الطبيعية والنفسية والاجتماعية .

إن كل قياس عددي يعبر عن موضع تركز التوزيع لمجموعة من البيانات يدعى " بمقياس النزعة المركزية " لهذه البيانات . إن وجود قيمة متوسطة تنزع إليها معظم القيم خاصية مفيدة في نواحي متعددة ، أبسطها أنه يمكن التعرف على مجموعة من البيانات من خلال قيمة واحدة هي القيمة المتوسطة . إن أحد أهم المقاييس المفيدة والأكثر استخداماً هو المعدل الوسطي (Average) ، وهو ما ندعوه غالباً بالوسط الحسابي (Arithmetic Mean) ، كذلك هناك الوسيط (Median) ، والمنوال (Mode) ولكل من هذه القيم المتوسطة معناه واستخداماته وطرق حسابه ، كما أنها تتساوى في بعض التوزيعات التكرارية وتختلف في توزيعات أخرى واختلاف قيمها أكثر توارداً من تساويها ولكن أهم تلك المتوسطات هو الوسط الحسابي .

4.2 المتوسط الحسابي (Arithmetic Mean)

يعتبر المتوسط الحسابي من أهم مقاييس النزعة المركزية بالرغم من وجود العديد من المتوسطات التي سوف نقوم بدراستها ضمن هذا الباب ، مثل الوسط التوافقي والوسط الهندسي وغيرها ويرمز له بالرمز \bar{x} ويمكن تعريفه بالشكل التالي :

إذا كانت لدينا مجموعة من القياسات وهي $X_1, X_2, X_3, \dots, X_n$ والتي ليس بالضرورة أن تكون مختلفة ، فإن الوسط الحسابي لهذه القياسات يعرف كما يلي :

$$\frac{\text{مجموع القيم}}{\text{عدد القيم}} = (\text{Arithmetic Mean})$$

أو بمعنى آخر فإن الوسط الحسابي يحسب من المعادلة الرياضية التالية :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n 1} \dots\dots\dots (1-2)$$

حيث أن :

\bar{x} : يمثل الوسط الحسابي .

i : يمثل رتبة القيمة الموجودة ضمن مجموعة القياسات التي تمثل مجتمعاً محدوداً حجمه n أما الرمز \sum فيرمز لمجموع القيم .

ويمكن كتابة القيمة الموجودة في البسط $\sum_{i=1}^n x_i$ بالرمز $\sum x$ حتى لا يكون هناك أي نوع من الالتباس . إن المعادلة (1-2) يمكن كتابتها أيضاً على الصورة التالية :

$$\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \dots\dots\dots (2-2)$$

5.2 طرق حساب المتوسط الحسابي

(Methods of Calculating Arithmetic Mean)

1- في حالة البيانات المفردة بدون تكرار

(For Data Set without Frequency)

إذا كان المطلوب إيجاد الوسط الحسابي لمجموعة من القياسات عددها n فإنه يمكن استخدام المعادلة (2-2) لذلك الغرض والمثال (1-2) يوضح لن ذلك .

مثال (1-2)

أوجد الوسط الحسابي للقيم التالية: 120 , 100 , 75 , 60 , 125

الحل :

نلاحظ أن عدد القيم هو 5 نستخدم المعادلة (1-2) أو (2-2) حيث :

$$\begin{aligned}\bar{x} &= \frac{\sum x_i}{n} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \\ &= \frac{125 + 60 + 75 + 100 + 120}{5} = \frac{480}{5} = 96\end{aligned}$$

مثال (2-2)

في خمس مباريات أشترك فيها فريق مدينة ما في كرة السلة قام بتسجيل النقاط التالية في تلك المباريات وكانت بالشكل التالي :

99 , 85 , 92 , 104 , 87

المطلوب :

أيجاد الوسط الحسابي للنقاط المسجلة لهذا الفريق .

الحل :

نلاحظ هنا أيضاً أن عدد المباريات هو خمسة أي أن $n = 5$ لذلك فإن :

$$\begin{aligned}\bar{x} &= \frac{\sum x_i}{n} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \\ &= \frac{87 + 104 + 92 + 85 + 99}{5} = \frac{467}{5} = 93.4\end{aligned}$$

مثال (2-3)

أخذت 6 سجائر كعينات من علب مختلفة من الدخان وسجلت كمية النيكوتين في كل منها بالمليجرام فكانت كما يلي :

21.2 , 16 , 10 , 15.7 , 18.1 , 12.3

المطلوب :

أحسب متوسط كمية النيكوتين في هذه العينة .

الحل :

نلاحظ هنا أن $n = 6$ وعلى هذا الأساس فإن الوسط الحسابي للعينة هو :

$$\begin{aligned}\bar{x} &= \frac{\sum x_i}{n} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \\ &= \frac{12.3 + 18.1 + 15.7 + 10 + 16 + 21.2}{6} = \frac{93.3}{6} = 15.55 \text{ mg}\end{aligned}$$

2- في حالة البيانات المفردة مع وجود تكرار (For Data Set with Frequency)

إذا كان المطلوب إيجاد المتوسط الحسابي لقيم مختلفة متكررة مثل القيم المبينة في الجدول (2-2) :

جدول (2-2)

x_n	x_3	x_2	x_1	القيم (x_i)
f_n	f_3	f_2	f_1	التكرار (f_i)

فإن المتوسط الحسابي يحسب من المعادلة التالية :

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{f_1 + f_2 + f_3 + \dots + f_n}$$

$$\therefore \bar{x} = \frac{\sum x_i f_i}{n} \quad \dots\dots\dots (3-2)$$

حيث أن :

f_i - يمثل تكرار القيمة في البيانات و n هو مجموع التكرارات .

مثال (4-2)

أوجد الوسط الحسابي لنتيجة الامتحان النهائي لعدد 100 من طلبة الفصل الأول في مادة الرياضيات في أحد المعاهد المهنية العليا ، والموضحة في الجدول (2-3) علماً أن النهاية العظمى لدرجة الامتحان هي 10 .

جدول (3-2)

10	9	8	7	6	5	4	3	2	1	0	الدرجة (x_i)
1	3	6	14	22	24	12	8	5	3	2	التكرار (f_i)

الحل:

نستخدم المعادلة رقم (3-2) ، لأن القيم هنا مكررة ولذلك فإن الوسط الحسابي لهذه القيم يحسب بالشكل التالي :

$$\bar{x} = \frac{\sum x_i f_i}{n}$$

$$\sum x_i f_i = (0)(2) + (1)(3) + (2)(5) + (3)(8) + (4)(12) + (5)(24) \\ + (6)(22) + (7)(14) + (8)(6) + (9)(3) + (10)(1) = 520$$

$$n = \sum f_i = 2 + 3 + 5 + 8 + 12 + 24 + 22 + 14 + 6 + 3 + 1 = 100$$

$$\therefore \bar{x} = \frac{520}{100} = 5.2$$

مثال (5-2)

من إحصاء عدد الناجحين في شهادة الثانوية العامة للفرع العلمي لعام (1980-1981) الحاصلين على معدل دون 50% ، بعد استثناء مادة التربية الإسلامية في مدينة عمان كما تم إدراج مجموعهم وأعدادهم في الجدول (4-2) :

جدول (4-2)

102	101	100	99	98	97	96	95	94	93	92	المجموع (Σx_i)
85	81	99	87	51	62	58	23	12	7	2	تعدد التكرار (f_i)

أوجد الوسط الحسابي لهذه العينة.

الحل :

$$\bar{x} = \frac{\sum x_i f_i}{n}$$

$$\sum x_i f_i = (92)(2) + (93)(7) + (94)(12) + (95)(23) + (96)(58) + (97)(62) \\ + (98)(51) + (99)(87) + (100)(99) + (101)(81) + (102)(85)$$

$$\therefore \sum x_i f_i = 56092$$

$$n = \sum f_i = 2 + 7 + 12 + 23 + 58 + 62 + 51 + 87 + 99 + 81 + 85 = 567$$

$$\therefore \bar{x} = \frac{56092}{567} = 98.9$$

مثال (6-2)

في إحدى التجارب المعملية تم الحصول على قياسات لدرجة حرارة ثرمومتر وقد أدرجت تلك القياسات وتكراراتها في الجدول (5-2) :

جدول (5-2)

5	8	6	2	القيمة (القياس) x_i
3	2	4	1	التكرار (f_i)

المطلوب : إيجاد الوسط الحسابي لهذه العينة .

الحل :

$$\bar{x} = \frac{\sum x_i f_i}{n}$$

$$\sum x_i f_i = (1)(2) + (6)(4) + (8)(2) + (5)(3)$$

$$\therefore \sum x_i f_i = 57$$

$$n = \sum f_i = 1 + 4 + 2 + 3 = 10$$

$$\therefore \bar{x} = \frac{57}{10} = 5.7$$

3- في حالة التوزيعات التكرارية ذات الفترات
(For Frequency Distributions with Intervals)

في هذه الحالة يمكن استخدام العلاقة (3-2) على أن نعتبر مراكز الفترات هي القيم المتكررة . وفي هذه الحالة فإن العلاقة (3-2) يمكن كتابتها على النحو التالي :

$$\bar{x} = \frac{\sum_{i=1}^n \bar{x}_i f_i}{\sum_{i=1}^n f_i} = \frac{\bar{x}_1 f_1 + \bar{x}_2 f_2 + \bar{x}_3 f_3 + \dots + \bar{x}_n f_n}{f_1 + f_2 + f_3 + \dots + f_n}$$
$$\therefore \bar{x} = \frac{\sum_{i=1}^n \bar{x}_i f_i}{n} \dots\dots\dots(4-2)$$

ومن المهم أن نذكر هنا أن ظهور الآلات الحاسبة الحديثة وأجهزة الحاسوب وسهولة استخدامها قد جعل من حساب الوسط الحسابي عملية سهلة لا تستغرق الكثير من الوقت .

فإذا أردنا مثلاً معرفة الوسط الحسابي للتوزيع التكراري الموضح في الجدول (3-1) أي حساب المتوسط الحسابي لدرجات اختبار الذكاء للخمسين مهندس ، نجد في الفئة الأولى ثلاثة مهندسين حصلوا على درجات تتراوح بين 90 وأقل من 100 ، وبافتراض أن المهندسين في كل فئة حصلوا على درجات متساوية وأن كلاً منها يساوي مركز الفئة (95) ، فإن مجموع الدرجات التي حصل عليها مهندسو الفئة الأولى يكون (95(مركز الفئة) \times 3 (تكرار الفئة)) ، ويكون مجموع الدرجات التي حصل عليها مهندسو الفئة الثانية (105 \times 14).... وهكذا و يكون المجموع الكلي للدرجات التي حصل عليها كل المهندسين مساوياً $(\sum \bar{x}_i f_i)$ ، وبقسمة هذا المجموع على عدد المهندسين $(\sum f_i)$ ، نحصل على الوسط الحسابي لدرجة المهندس في الاختبار كما يوضح الجدول (6-2) ذلك .

جدول (6-2)

إيجاد الوسط الحسابي لدرجة المهندس في اختبار الذكاء

ت	الفئة (x_i)	التكرار (f_i)	مركز الفئة (\bar{x}_i)	مجموع تكرار الفئة ($\bar{x}_i f_i$)
1	- 90	3	95	$285 = 3 \times 95$
2	- 100	14	105	$1470 = 14 \times 105$
3	- 110	16	115	$1840 = 16 \times 115$
4	- 120	11	125	$1375 = 11 \times 125$
5	- 130	4	135	$540 = 4 \times 135$
6	150 - 140	2	145	$290 = 2 \times 145$
	المجموع Total	50	————	5800

وبواسطة المعادلة (2-4) يمكن حساب الوسط الحسابي لدرجات اختبار
النكاه للخمسين مهندس بالشكل التالي :

$$\bar{x} = \frac{\sum_{i=1}^n \bar{x}_i f_i}{n} = \frac{5800}{50} = 116$$

ونظراً لكثرة العمليات الحسابية التي تتضمنها هذه الطريقة حيث نقوم بضرب أرقام التكرارات في مراكز الفئات التي قد تحتوي على كسور لذلك يفضل استخدام طرق أخرى مختصرة لتبسيط العمليات الحسابية .

مثال (2-7)

من إحصاء عدد الناجحين في شهادة الثانوية العامة الفرع العلمي العام
للعام الدراسي (1992-1993) في مدينة عمان وجد أن المجاميع كما يبين
الجدول (2-7) :

جدول (2-7)

200 -	190 -	180 -	170 -	160 -	150 -	140 -	130 -	120 -	110 -	الفئة
1	2	5	9	21	37	102	135	252	415	التكرار

المطلوب :

أوجد الوسط الحسابي لهذه العينة .

الحل

نرتب الجدول بالشكل الذي يبينه الجدول (2-8) :

جدول (8-2)

ن	حدود الفئة x_i	مركز الفئة \bar{x}_i	التكرار f_i	مركز الفئة x التكرار $\bar{x}_i f_i$
1	110 -	115	415	47725
2	120 -	125	252	31500
3	130 -	135	135	18225
4	140 -	145	102	14790
5	150 -	155	37	5735
6	160 -	165	21	3465
7	170 -	175	9	1575
8	180 -	185	5	925
9	190 -	195	2	390
10	200 -	205	1	205
المجموع Total				124535
				979

$$\therefore \bar{x} = \frac{\sum_{i=1}^n \bar{x}_i f_i}{n} = \frac{124535}{979} = 127.2$$

مثال (8-2)

الجدول (9-2) يمثل توزيع دخل 100 منتج في أحد المصانع والمطلوب حساب الوسط الحسابي لهذا الدخل .

جدول (9-2)

فترات الأجور	40 -	46 -	52 -	58 -	64 -	70 -	76 -	82 -	88 -	94 - 100
التكرار	2	10	12	15	20	17	11	8	4	1

الحل :

في هذا المثال نعتبر مراكز الفئات هي الدخل الذي نريد إيجاد وسطه الحسابي والخطوات موضحة في الجدول أدناه ومن الجدول نجد أن :

$$\sum f_i = n = 100 \text{ , } \sum x_i f_i = 6742$$

ولذلك نحسب الوسط الحسابي كما يبين الجدول (10-2) :

جدول (10-2)

ت	فئات الدخل x_i	عدد المنتجين (التكرار)	مراكز الفئات	المركز x التكرار
1	-40	2	43	86
2	-46	10	49	490
3	-52	12	55	660
4	-58	15	61	915
5	-62	20	67	1340
6	-68	17	73	1241
7	-76	11	79	869
8	-82	8	85	680
9	-88	4	91	364
10	100 -94	1	97	97
المجموع Total		100	—	6742

$$\therefore \bar{x} = \frac{\sum x_i f_i}{n} = \frac{6742}{100} = 67.42$$

6.2 حساب الوسط الحسابي بطريقة الوسط الفرضي (Finding Arithmetic Mean Using Assumed Mean)

تتضمن هذه الطريقة اختيار وسطاً فرضياً وحساب الانحرافات عن هذا الوسط الفرضي ، فإذا كان مجموع هذه الانحرافات صفراً ، فإن ذلك يدل على أن الوسط الفرضي الذي اخترناه هو نفسه المتوسط الحسابي ، وإذا كان مجموع الانحرافات موجباً كان المتوسط الحسابي أكبر من الوسط الفرضي ، وإذا كان مجموع الانحرافات سالباً كان المتوسط الحسابي أصغر من الوسط الفرضي وعلى هذا الأساس فإن المتوسط الحسابي يعرف من المعادلة التالية :

المتوسط الحسابي الأصلي = الوسط الفرضي + متوسط الانحرافات

وإذا رمزنا للوسط الفرضي بالرمز a وإلى مجموع الانحرافات عن الوسط الفرضي بالرمز d حيث إن :

$$d = \frac{\sum_{i=1}^n (x_i - a)}{n}$$

فإنه يمكن حساب الوسط الحسابي كما يلي:

$$\bar{x} = a + d = a + \frac{\sum (x_i - a)}{n} \dots\dots\dots(5-2)$$

ولتوضيح كيفية حساب الوسط الحسابي بطريقة الوسط الفرضي نقوم بدراسة المثال (9-2) والذي يوضح لنا ذلك .

مثال (2 - 9)

باستخدام طريقة الوسط الفرضي ، أوجد الوسط الحسابي للبيانات التالية :

30 , 15 , 20 , 12 , 5 , 8

الحل:

نلاحظ أن القياسات غير مرتبة لذلك نقوم بترتيبها تصاعدياً أو تنازلياً بالشكل التالي :

30 , 20 , 15 , 12 , 8 , 5

ثم نختار وسطاً فرضياً من بين القيم وليكن العدد 12 ثم نحسب مجموع الانحرافات عن الوسط الفرضي بالشكل الموضح في الجدول (2-11) :

جدول (2-11)

ت	القيمة	الوسط الفرضي	قيمة الانحراف عن الوسط الفرضي
1	5	12	7 - - 12 - 5
2	8	12	4 - - 12 - 8
3	12	12	0 - 12 - 12
4	15	12	3 = 12 - 15
5	20	12	8 = 12 - 20
6	30	12	18 = 12 - 30
المجموع Total		—	مجموع الانحرافات = 18

وسنلاحظ أن :

مجموع الانحرافات = 18 عدد القيم أو الحالات = 6

ولذلك فإن متوسط الانحرافات تساوي $18/6=3$ ، وعلى هذا الأساس فإن الوسط الحسابي الأصلي يساوي $12 + 3 = 15$ ، ويمكن التحقق من ذلك بتطبيق التعريف مباشرة حيث نجد أن :

$$\bar{x} = \frac{5+8+12+15+20+30}{6} = \frac{90}{6} = 15$$

مثال (2 - 10)

أوجد الوسط الحسابي بطريقة الوسط الفرضي للبيانات التالية :

2 , 9 , 8 , 13 , 10 , 12

الحل:

نرتب القيم تصاعدياً فتكون على النحو التالي:

13 , 12 , 10 , 9 , 8 , 2

ثم نختار الوسط الفرضي بطريقة بحيث تكون الانحرافات السالبة عن يمينه تساوي الانحرافات الموجبة عن يساره ، أي نختار أما العدد 9 أو العدد 10 ولنأخذ العدد 10 مثلاً كوسط فرضي . نرتب القيم كما هو مبين في الجدول (2-12) ، ونلاحظ بأن مجموع الانحرافات يساوي (- 6) وعدد القيم أو الحالات تساوي 6 أيضاً ، ولذلك فإن متوسط الانحرافات تساوي $-6/6 = -1$ وعلى هذا الأساس فإن الوسط الحسابي الأصلي سيساوي $10 + (-1) = 9$ ويمكن التحقق من ذلك بتطبيق التعريف مباشرة حيث نجد أن :

جدول (12-2)

ت	القيمة	الوسط الفرضي	قيمة الانحراف عن الوسط الفرضي
1	2	10	8 - = 10 - 2
2	8	10	2 - = 10 - 8
3	9	10	1 - = 10 - 9
4	10	10	0 = 10 - 10
5	12	10	2 = 10 - 12
6	13	10	3 = 10 - 13
المجموع Total		—	مجموع الانحرافات - = 6

$$\therefore \bar{x} = \frac{2+8+9+10+12+13}{6} = \frac{54}{6} = 9$$

مثال (2 - 11)

أوجد الوسط الحسابي بطريقة الوسط الفرضي للقياسات التالية :

25 , 20 , 15 , 10 , 6 , 2

الحل :

نلاحظ أن القياسات مرتبة تصاعدياً ولذلك نختار العدد 15 كوسط فرضي . ونرتب القيم في الجدول (13-2) .

جدول (13-2)

ت	القيمة	الوسط الفرضي	قيمة الانحراف عن الوسط الفرضي
1	2	15	13 - = 15 - 2
2	6	15	9 - = 15 - 6
3	10	15	5 - = 15 - 10
4	15	15	0 = 15 - 15
5	20	15	5 = 15 - 20
6	25	15	10 = 15 - 25
المجموع Total		—	مجموع الانحرافات - = 12

$$\therefore \bar{x} = 15 + \frac{-12}{6} = 15 - 2 = 13$$

أما عند حساب المتوسط الحسابي في حالة البيانات المبوبة أي البيانات المرتبة في جداول التوزيعات التكرارية والمكونة من فئات فإنه يجب اتباع الخطوات التالية :

- 1- نجد مركز الفئة \bar{x}_i .
- 2- نختار وسطاً فرضياً (a) من بين قيم \bar{x}_i ، ومن المفضل أن تكون القيمة الوسطية للفئات .
- 3- نطرح قيم \bar{x}_i من الوسط الفرضي ونكون بذلك قد وجدنا الانحرافات عن الوسط الفرضي (d) .
- 4- نضرب قيمة الانحراف عن الوسط الفرضي (d) في تكرار الفئة (f) فنحصل على مجموع الانحرافات لتلك الفئة .
- 5- نجمع الانحرافات لكل الفئات ونقسمها على مجموع التكرارات فنحصل على معدل الانحرافات للفئات .
- 6- نضيف معدل الانحرافات إلى الوسط الفرضي فنحصل على الوسط الحسابي وكما موضح في المعادلة التالية :

$$\bar{x} = a + \frac{\sum d_i f_i}{\sum f_i} = a + \frac{\sum d_i f_i}{n} \dots\dots\dots (6 - 2)$$

حيث أن d_i يمثل انحراف قيم مراكز الفئات عن الوسط الفرضي (a) والذي يتم اختياره من بين الفئات أما f_i فهو تكرار الفئة. ويراعى عند اختيار الوسط الفرضي أن يكون في منتصف الجدول ويستحسن أن يكون أمام أكبر تكرار . ويوضح الجدول (2-14) كيفية إيجاد الوسط الحسابي بطريقة الوسط الفرضي لاختبار درجات الذكاء للخمسين مهندس .

ولتوضيح كيفية ترتيب الجدول لإيجاد الوسط الحسابي باستخدام طريقة الانحرافات أي الوسط الغرضي نتبع ما يلي :

جدول (14-2)

الفئات	تكرار الفئة f_i	مركز الفئة \bar{x}_i	قيمة الانحراف عن الوسط الغرضي d_i	مجموع الانحرافات $d_i f_i$
90 -	3	95	20 -	60 -
100 -	14	105	10 -	140 -
110 -	16	115	0	0
120 -	11	125	10	110
130 -	4	135	20	80
140 - 150	2	145	30	60
المجموع	50	—	—	50

- 1- نعين مراكز الفئات كما موضح في الجدول .
- 2- نختار وسطاً فرضياً من بين مراكز الفئات وهو في هذه الحالة 115 وكما تم توضيح ذلك ، أي المركز المقابل لأكبر تكرار ونحسب انحرافات مراكز الفئات عنه .
- 3- نضرب تكرار كل فئة (f_i) في انحراف هذه الفئة (d_i) .
- 4- نجمع حاصل الضرب فنحصل على $\sum f_i d_i$ ، ثم نوجد الوسط الحسابي بواسطة المعادلة (6-2) . وبإتباع هذه الخطوات نجد أن الوسط الحسابي هو :

$$\bar{x} = a + \frac{\sum d_i f_i}{\sum f_i} = 115 + \frac{50}{50} = 115 + 1 = 116$$

مثال (2- 12)

أوجد الوسط الحسابي بطريقة الوسط الفرضي الانحرافات لدخول 100 منتج والمعطاة في المثال (2 - 8) .

الحل:

نرتب الجدول القيم كما يبين الجدول (2-15) حيث :

جدول (2-15)

القيمات x_i	تكرار الفئة f_i	مركز الفئة \bar{x}_i	قيمة الانحراف عن الوسط الفرضي a	مجموع الانحرافات $d_i f_i$
- 40	2	43	- 24	- 48
- 46	10	49	- 18	- 180
- 52	12	55	- 12	- 144
- 58	15	61	- 6	- 90
- 64	20	67	0	0
- 70	17	73	6	102
- 76	11	79	12	132
- 82	8	85	18	144
- 88	4	91	24	96
100 - 94	1	97	30	30
المجموع	100	—	—	36

$$\bar{x} = a + \frac{\sum d_i f_i}{\sum f_i} = 67 + \frac{36}{100} = 67 + 0.36 = 67.36$$

7.2 حساب الوسط الحسابي بطريقة الانحرافات المختصرة (Calculation of Arithmetic Mean Using Shortcut Deviations)

لزيادة اختصار العمليات الحسابية نلاحظ أنه في حالة الجداول المنتظمة أي عند تساوي أطوال الفئات يمكن قسمة كل من الانحرافات (d) على طول الفئة (L) لنحصل بذلك على الانحرافات المختصرة أي أن :

$$\bar{d} = \frac{d}{L} \dots \dots \dots (7-2)$$

وبذلك نحصل على الوسط الحسابي من العلاقة التالية :

$$\bar{x} = a + \frac{\sum \bar{d}_i f_i}{\sum f_i} L = a + \frac{\sum \bar{d}_i f_i}{n} L \dots \dots \dots (8-2)$$

حيث أن \bar{d} هي الانحرافات المختصرة ، L هو طول الفئة ويلاحظ أن إتباع هذه الطريقة يسهل من العمليات الحسابية كثيراً ويكشف الأخطاء في حساب الانحرافات ، حيث أن الانحرافات المختصرة للفئات التي تلي فئة الوسط الفرضي تكون دائماً 1 ، 2 ، 3 والتي تسبق فئة الوسط الفرضي تكون - 1 ، - 2 ، - 3 .. الخ .

ويبين الجدول (16-2) كيفية إيجاد الوسط الحسابي بطريقة الانحرافات المختصرة عن الوسط الفرضي لدرجات اختبار الذكاء للخمسين مهندس ، حيث تراعى نفس الخطوات السابقة مع مراعاة وضع :

$$\bar{d} = \frac{d}{L}$$

حيث أن L تمثل طول الفترة ويلاحظ أن النتيجة النهائية هي نفسها التي تم حسابها بواسطة الطريقة السابقة .

جدول (16-2)

إيجاد الوسط الحسابي باستخدام طريقة الانحرافات المختصرة لدرجات اختبار الذكاء للخمسين مهندس

الفئات	تكرار الفئة f_i	مركز الفئة \bar{x}_i	قيمة الانحراف عن الوسط الفرضي d_i	الانحراف المختصر عن الوسط الفرضي \bar{d}_i	مجموع الانحرافات المختصرة $f_i \bar{d}_i$
90 -	3	95	20 -	2 -	6 -
100 -	14	105	10 -	1 -	14 -
110 -	16	115	0	0	0
120 -	11	125	10	1	11
130 -	4	135	20	2	8
140 - 150	2	145	30	3	6
المجموع	50	—	—	—	5

$$\therefore \bar{x} = a + \frac{\sum \bar{d}_i f_i}{n} = 115 + \left(\frac{5}{50} \times 10 \right) = 115 + 1 = 116$$

وهي نفس النتيجة التي حصلنا عليها من قبل . ويعتبر المتوسط الحسابي من أكثر المتوسطات استخداماً في علم الإحصاء ، وذلك لأنه يستخدم كل القيم المعطاة في البيانات ، كما أنه أكثر المتوسطات ثباتاً من عينة إلى أخرى للتعبير عن خواص مجتمع معين . هذا بالإضافة إلى الدقة في طريقة حسابه غير أنه في بعض الأحيان لا يكون المتوسط معبراً تماماً عن الصورة الحقيقية للبيانات ، ففي حالة وجود بعض القيم المتطرفة في الصغر مثلاً فإننا نجد أن الوسط ينجذب ناحية القيم الصغرى رغم أن معظم

الحالات تكون ذات قيم كبرى ، وقد يحدث وضع شاذ مماثل في حالة تطرف عدد قليل من البيانات الكبيرة القيمة فنجد أن الوسط ينجذب إلى القيم الكبرى رغم أن معظم الحالات تكون قيماً صغيرة . ولعلك سمعت أو قرأت في بعض الأحيان عن مواقف بساء فيها استخدام المتوسط للتعبير عن مجموعة من البيانات فمثلاً قد نجد عشرة أشخاص في إحدى المنشآت مرتباتهم كما يبين الجدول (17-2) :

جدول (17-2)

الوظيفة	العدد	المرتّب (بالدينار)
المدير	1	800
الكاتب	1	80
المنتجون	8	60

أن المتوسط الحسابي لهذه المرتبات هو 136 دينار . فإذا أدعى المسؤول عن العمل بأن متوسط المرتبات هو 120 دينار فإن هذا الإدعاء لا يمثل الصورة الحقيقية لدخل المنتجين ، والتي هي أقل من هذا المبلغ ولكن لأن مرتب المدير قد رفع المتوسط الحسابي ، ومن هنا يتضح أن المتوسط الحسابي لن يعطي الصورة الحقيقية لمتوسط المرتبات في هذه المنشأة .

مثال (2 - 13)

أوجد الوسط الحسابي بطريقة الانحرافات المختصرة لدخل 100 منتج والمعطاة في المثال (2 - 8) .

الحل :

نرتب القيم في الجدول (17-2) حيث :

جدول (17-2)

الفئات x_i	تكرار الفئة f_i	مركز الفئة \bar{x}_i	قيمة الانحراف عن الوسط الفرضي به	معدل الانحرافات المختصرة \bar{d}_i	مجموع الانحرافات المختصرة $\bar{d}_i f_i$
- 40	2	43	24 -	4 -	8 -
- 46	10	49	18 -	3 -	30
- 52	12	55	12 -	2 -	24 -
- 58	15	61	6 -	1 -	15 -
- 64	20	67	0	0	0
- 70	17	73	6	1	17
- 76	11	79	12	2	22
- 82	8	85	18	3	24
- 88	4	91	24	4	16
100 - 94	1	97	30	5	5
المجموع	100	—	—	—	7

$$\therefore \bar{x} = a + \frac{\sum \bar{d}_i f_i}{n} = 67 + \left(\frac{7}{100} \times 6 \right) = 67 + 0.42 = 67.42$$

نلاحظ أن قيمة المتوسط الحسابي المحسوبة بهذه الطريقة هي نفس القيمة المحسوبة في المثال (2 - 8) .

8.2 خواص المتوسط الحسابي (Arithmetic Mean Properties)

من أهم خواص الوسط الحسابي النقاط التالية :

- 1- سهولة حسابه وخضوعه للعمليات الجبرية .
- 2- مجموع انحرافات القيم عن وسطها الحسابي يساوي صفراً ، ويمكن اثبات ذلك جبرياً بالشكل التالي :

$$d_1 = x_1 - \bar{x}$$

$$d_2 = x_2 - \bar{x}$$

$$\vdots \quad \quad \quad \vdots$$

$$d_n = x_n - \bar{x}$$

وبالجمع نجد أن :

$$\sum d_i = \sum x_i - n \bar{x}$$

أي أن :

$$\therefore \sum d_i = \sum x_i - n \left(\frac{\sum x_i}{n} \right) = 0$$

- 3- مجموع انحرافات مربع القيم عن متوسطها الحسابي يقل عن مجموع انحرافات مربعات القيم عن أية قيمة أخرى وسنثبت ذلك فيما بعد .

- 4- إذا كان لدينا عدداً من القيم لمتغيرين وهما x و y فإن المتوسط الحسابي لمجموع قيم المتغيرين يساوي مجموع الوسطين الحسابيين للمتغيرين .

فإذا كان :

$$z_1 = x_1 + y_1$$

$$z_2 = x_2 + y_2$$

$$\vdots \quad \quad \quad \vdots$$

$$z_n = x_n + y_n$$

وبالجمع نجد أن :

$$\sum z = \sum x + \sum y$$

وبالقسمة على عدد القيم n نجد أن :

$$\bar{z} = \bar{x} + \bar{y}$$

ويمكن تعميم هذه النتيجة في حالة أي عدد من المتغيرات فإذا كانت :

$$m = x + y - z$$

فإن :

$$\bar{m} = \bar{x} + \bar{y} - \bar{z}$$

5- إذا احتوت مجموعة من المفردات على بعض القيم المتطرفة الكبيرة جداً أو الصغيرة جداً ، فإن المتوسط الحسابي للمجموعة يكون مضللاً لأنه يتأثر بهذه القيم بالرغم من قلّتها وفي مثل هذه الحالة يفضل الاعتماد على مقياس آخر من مقاييس المتوسطات .

6- لا يمكن حساب الوسط الحسابي من الجداول المفتوحة .

7- لا يمكن إيجاد الوسط الحسابي بالطرق البيانية .

وفي بعض الأحيان قد لا نستطيع حساب المتوسط حينما يكون لدينا توزيع تكراري في فترات مفتوحة سواء كانت من أدنى أو من أعلى حيث لا نتمكن من معرفة مركز الفترة المفتوحة ولمثل هذه الأسباب قد نلجأ إلى المؤشر الثاني لمقاييس النزعة المركزية المسمى بالوسيط .

1.8.2 المتوسط الحسابي المرجح

(Predominant Arithmetic Mean)

إذا كانت القيم المشاهدة ليس لها نفس الأهمية أو الوزن ، فعندئذٍ لحساب المتوسط الحسابي لتلك المجموعة يجب أن لا نعامل جميع القيم نفس المعاملة بل نكرر كل قيمة عدد من المرات حسب أهميتها أو وزنها ، أي نرجح كل قيمة بوزنها وعليه يطلق على المتوسط الحسابي في هذه الحالة المتوسط الحسابي المرجح أو الموزون .

فإذا فرضنا أن لدينا القيم التالية : $x_1, x_2, x_3, \dots, x_n$ وكانت أهمية كل قيمة متناسبة مع الأوزان : $w_1, w_2, w_3, \dots, w_n$ فيحسب المتوسط الحسابي المرجح من العلاقة التالية :

$$\tilde{x} = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3 + \dots + x_n w_n}{w_1 + w_2 + w_3 + \dots + w_n} = \frac{\sum_{i=1}^n x_i w_i}{\sum_{i=1}^n w_i} \dots\dots\dots (9-2)$$

7- لا يمكن إيجاد الوسط الحسابي بالطرق البيانية .

وفي بعض الأحيان قد لا نستطيع حساب المتوسط حينما يكون لدينا توزيع تكراري في فترات مفتوحة سواء كانت من أدنى أو من أعلى حيث لا نتمكن من معرفة مركز الفترة المفتوحة ولمثل هذه الأسباب قد نلجأ إلى المؤشر الثاني لمقاييس النزعة المركزية المسمى بالوسيط .

1.8.2 المتوسط الحسابي المرجح

(Predominant Arithmetic Mean)

إذا كانت القيم المشاهدة ليس لها نفس الأهمية أو الوزن ، فعندئذٍ لحساب المتوسط الحسابي لتلك المجموعة يجب أن لا نعامل جميع القيم نفس المعاملة بل نكرر كل قيمة عدد من المرات حسب أهميتها أو وزنها ، أي نرجح كل قيمة بوزنها وعليه يطلق على المتوسط الحسابي في هذه الحالة المتوسط الحسابي المرجح أو الموزون .

فإذا فرضنا أن لدينا القيم التالية : $x_1, x_2, x_3, \dots, x_n$ وكانت أهمية كل قيمة متناسبة مع الأوزان : $w_1, w_2, w_3, \dots, w_n$ فيحسب المتوسط الحسابي المرجح من العلاقة التالية :

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3 + \dots + x_n w_n}{w_1 + w_2 + w_3 + \dots + w_n} = \frac{\sum_{i=1}^n x_i w_i}{\sum_{i=1}^n w_i} \dots\dots\dots(9-2)$$

ونلاحظ أن صيغة المتوسط الحسابي المرجح مماثلة لصيغة المتوسط الحسابي في حالة البيانات المبوبة في جدول تكراري لأن التكرار يعبر عن وزن أو أهمية القيمة .

فعلى سبيل المثال إذا كان لدينا 3 سلع ، وكان سعر الوحدة من السلعة الأولى 26 ديناراً وسعر الوحدة من السلعة الثانية 20 ديناراً وسعر الوحدة من السلعة الثالثة 6 دنائير ، وكانت أهمية السلعة الثانية ضعف أهمية السلعة الأولى وكانت أهمية السلعة الثالثة خمسة أمثال السلعة الأولى فإن الوسط الحسابي المرجح يحسب كما هو مبين في الجدول (18-2) :

جدول (18-2)

المفردة	السلعة الأولى	السلعة الثانية	السلعة الثالثة
سعر الوحدة (x_i)	26	20	6
الأهمية (w_i)	1	2	5

وبتطبيق الصيغة الأولى للمتوسط الحسابي المرجح نحصل على :

$$\bar{x} = \frac{x_1 w_1 + x_2 w_2 + \dots + x_n w_n}{w_1 + w_2 + \dots + w_n} = \frac{(26)(1) + (20)(2) + (6)(5)}{1 + 2 + 5} = 12$$

9.2 الوسط (The Median)

يعرف الوسيط بأنه القيمة التي تتوسط البيانات أي القيمة التي تقع في المنتصف ، فإذا أردنا إيجاد الوسيط لمجموعة من المفردات ، فإننا نقوم بترتيب هذه المجموعة تصاعدياً أو تنازلياً ثم نبحث عن القيمة التي يسبقها

وبليها نفس العدد من القيم ، ويمكن تعريف الوسيط أيضا بأنه القيمة التي تقسم المجموعة إلى قسمين بحيث يكون عدد القيم الأصغر منه مساوياً لعدد القيم الأكبر منه .

1.9.2 خواص الوسيط (The Median Properties)

- 1- يصف البيانات بواقعية أكثر في حالة وجود القيم المتطرفة .
- 2- يمكن إيجاده من الجداول المفتوحة .
- 3- يمكن تحديده من الرسم البياني .
- 4- لا يتأثر بالقيم المتطرفة .
- 5- ليس له معنى في حالة عدم ترتيب البيانات تصاعدياً أو تنازلياً .

2.9.2 طرق حساب الوسيط (Methods for Calculating Median)

أولاً : في حالة البيانات المفردة نقوم لحساب الوسيط بإتباع الخطوات التالية :

- (a) نرتب البيانات تصاعدياً أو تنازلياً .
- (b) نحدد رتبة الوسيط فإذا كانت القيم فردية فإن رتبة الوسيط تساوي :

$$\frac{n + 1}{2}$$

- (c) أما إذا كانت القيم زوجية العدد فإن رتبة الوسيط هي :

$$\left(\frac{n}{2}, \frac{n}{2} + 1 \right)$$

مثال (2-14)

أوجد الوسيط للقيم التالية :

5 , 15 , 3 , 10 , 8 , 6 , 20

الحل :

نرتب أولاً القيم تصاعدياً بالشكل التالي:

20 , 15 , 10 , 8 , 6 , 5 , 3

وبما أن عدد القيم فردي وهو 7 لذلك فإن رتبة الوسيط تساوي :

$$.4 = \frac{8}{2} = \frac{1+7}{2}$$

أي إن رتبة الوسيط هي الرتبة الرابعة من اليمين أي العدد 8 ($M = 8$) .

مثال (2 - 15)

أوجد الوسيط للقيم التالية :

60 , 135 , 55 , 40 , 45 , 40 , 50

الحل :

نقوم بترتيب القيم تصاعدياً ويكون الوسيط مساوياً لقيمة الحد الأوسط في

الترتيب بالشكل التالي :

135	60	55	50	45	40	40
-----	----	----	----	----	----	----



أي أن الوسيط يساوي 50 . ونلاحظ أن الوسيط هنا هو الحد الرابع ، أي أن رتبة الوسيط هي 4 حيث عدد الحالات هو 7 كما في المثال السابق .

مثال (16-2)

أوجد الوسيط للبيانات التالية :

3 , 5 , 9 , 12 , 18 , 30 , 25 , 20 , 32 , 21

الحل :

نلاحظ أن عدد البيانات هنا هو 10 وهو رقم زوجي لذلك نقوم أولاً بترتيب البيانات تصاعدياً كما يلي :

35 , 32 , 30 , 25 , 21 , 18 , 12 , 9 , 5 , 3

إن رتبة الوسيط هي :

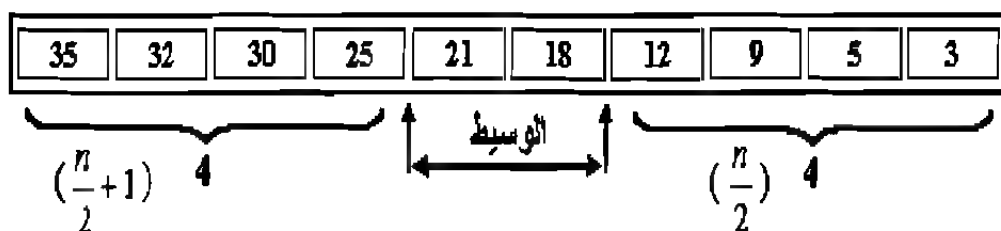
$$\frac{n}{2}, \frac{n}{2} + 1$$

أي أن :

$$\frac{10}{2}, \frac{10}{2} + 1 = 5.6$$

أي الترتيب الخامس والسادس وعلى هذا الأساس فإن العددين الذين بينهما الوسيط هما 21, 18 على التوالي ولذلك فإن الوسيط يحسب بالشكل التالي:

$$M = \frac{18 + 21}{2} = \frac{39}{2} = 19.5$$



ثانياً : في حالة البيانات المبوبة (In Case of Tabulated Data) أي البيانات المرتبة في جداول التوزيعات التكرارية فإنه يجب لحساب الوسيط اتباع الخطوات التالية :

- 1- نوجد أولاً جدول التكرار المتجمع الصاعد (أو النازل) للبيانات .
- 2- نحدد رتبة الوسيط والتي تساوي $\frac{n}{2}$.
- 3- نجد الفترة الوسيطة (Median Interval) التي تقع فيها رتبة الوسيط .
- 4- نطبق القانون حسب العلاقة (14-2) لإيجاد الوسيط حيث :

$$M = L + \frac{\frac{n}{2} - k_1}{k_2 - k_1} i \dots\dots\dots(14-2)$$

حيث أن :

- M : الوسيط .
- L : الحد الأدنى للفترة الوسيطة .
- $\frac{n}{2}$: رتبة الوسيط .
- k_1 : التكرار المتجمع الصاعد الذي يسبق الفترة الوسيطة .
- k_2 : التكرار المتجمع الصاعد الذي يلي الفئة الوسيطة .
- i : طول الفئة أو الفترة .

$$\text{الوسيط} = \text{الحد الأدنى للفئة الوسيطة} + \left[\frac{\text{ترتيب الوسيط} - \text{التكرار المتجمع الصاعد السابق لترتيب الوسيط}}{\text{التكرار المتجمع الصاعد اللاحق} - \text{التكرار المتجمع الصاعد السابق لترتيب الوسيط}} \right] \times \text{طول فئة الوسيط}$$

ولحساب الوسيط للتوزيع التكراري الخاص باختبار درجات الذكاء
للخمسين مهندس نكون جدول التكرار المتجمع الصاعد كما موضح في الجدول
(19-2) حيث :

جدول (19-2)

إيجاد الوسيط والربيعين للتوزيع التكراري الخاص
باختبار درجات الذكاء للخمسين مهندس

ت	الفئات	التكرار	حدود الفئات	التكرار المتجمع الصاعد
1	—	—	أقل من 90	صفر
2	90 -	3	أقل من 100	3
3	100 -	14	أقل من 110	17
<div style="text-align: center;"> الفئة الوسيطة </div>				
4	110 -	16	أقل من 120	33
5	120 -	11	أقل من 130	44
6	130 -	4	أقل من 140	48
7	140 - 150	2	أقل من 150	50

والخطوات هي كما يلي :

- 1- نحسب ترتيب الوسيط وهو في هذه الحالة يساوي $25 = \frac{50}{2} = \frac{n}{2}$.
- 2- نحدد الفئة الوسيطة كما يشير السهم فنجدها الفئة (110 وأقل من 120)
وهنا سنجد أن :

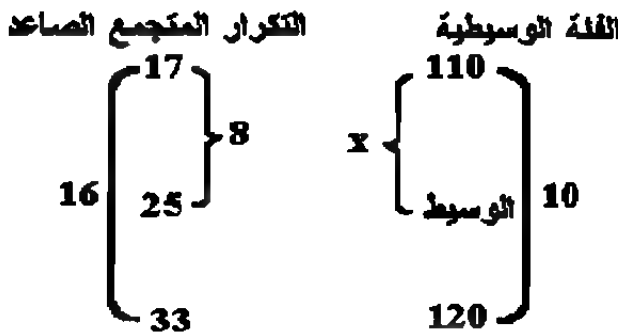
$$i = L_2 - L_1 = 120 - 110 = 10 \quad , \quad k_2 = 33 \quad , \quad k_1 = 17$$

3- بتطبيق المعادلة (2-14) نجد قيمة الوسيط بالشكل التالي :

$$M = L + \frac{\frac{n}{2} - k_1}{k_2 - k_1} i = 110 + 10 \left(\frac{25 - 17}{33 - 17} \right) = 110 + 10 \left(\frac{8}{16} \right)$$

$$\therefore M = 110 + \frac{10}{2} = 110 + 5 = 115$$

وهناك طريقة اخرى تتمثل بالشكل التالي :



1- بعد تكوين جدول التكرار المتجمع الصاعد نكون قوسين كما في الشكل أعلاه ، أحدهما بين الفئة الوسيطة والآخر يبين التكرار المتجمع الصاعد المناظر .

2- سوف نجد أن طول الفئة الوسيطة يساوي 10 وتكرار الفئة الوسيطة هو 16.

3- بحسابنا البعد بين ترتيب الوسيط (25) وبين التكرار المتجمع الصاعد السابق لفئة الوسيط وهو 17 سنجد بأن هذا البعد هو $8 = 17 - 25$.

4- إن هذا البعد يناظر بعد الوسيط عن الحد الأدنى للفئة الوسيطة والذي سنرمز له بالرمز x أي أن الوسيط $= 110 + x$ حيث أن x يمكن الحصول عليها من العلاقة التالية :

$$\frac{x}{10} = \frac{8}{16}$$

5- ولحساب قيمة x نجد أن :

$$80 = 16 x$$

أي أن :

$$x = 5$$

6- إذن الوسيط يساوي $110 + 5 = 115$ درجة .

وهناك معادلة أخرى يمكن حساب الوسيط بواسطتها وتتمثل بالشكل التالي :

$$\text{الوسيط} = \text{الحد الأدنى للفترة الوسيطة} + \left\{ \frac{\text{تركيب الوسيط - التكرار المتجمع السابق للفترة الوسيطة}}{\text{تكرار الفترة الوسيطة}} \right\} \times \text{طول الفترة الوسيطة}$$

ويمكن اختصار الصيغة أعلاه بالعلاقة (2-15) حيث :

$$M = L + \frac{\frac{n}{2} - k_1}{F} \times i \dots\dots\dots(15-2)$$

حيث يمثل F تكرار الفئة الوسيطة ولو أردنا تطبيق هذه القاعدة لإيجاد الوسيط لدرجات اختبار الذكاء للخمسين مهندس وبالعودة إلى الجدول (2-19) سنجد أن :

$$L = 110, k_1 = 17 \text{ وأيضاً } \frac{n}{2} = \frac{50}{2} = 25 \text{ و } F = 16 \text{ أما } i \text{ فتساوي } 10$$

وبتعويض هذه القيم في المعادلة (2 - 15) نحصل على:

$$M = L + \frac{\frac{n}{2} - k_1}{F} \times i = 110 + 10 \left(\frac{25 - 17}{16} \right) = 110 + \frac{(10)(8)}{(16)}$$

$$\therefore M = 110 + 5 = 115$$

وهي نفس القيمة السابقة .

مثال (2 - 17)

أوجد الوسيط لدخل الـ 100 منتج والمعطى في المثال (2 - 8) .

الحل:

نكون جدول التكرار المتجمع الصاعد بالصورة الموضحة في الجدول (20-2) وأتباع الخطوات التالية :

$$1- \text{ نحدد رتبة الوسيط وهي في هذا المثال تساوي } 50 = \frac{100}{2} = \frac{n}{2}$$

2- من جدول التكرار المتجمع الصاعد للفئات نلاحظ ما يلي :

$$L_1 = 58 , L_2 = 62 , i = 62 - 58 = 4 , k_1 = 39 , F = 15 , k_2 = 59$$

3- نعوض القيم أعلاه أما في المعادلة (2-14) أو المعادلة (2-15) فنحصل على قيمة الوسيط كما يلي :

$$M = L + \frac{\frac{n}{2} - k_1}{k_2 - k_1} i = 58 + 4 \left(\frac{50 - 39}{59 - 39} \right) = 58 + 4 \left(\frac{11}{20} \right)$$

$$\therefore M = 58 + \frac{11}{5} = 58 + 2.2 = 60.2$$

جدول (2-20)

ت	الفئات	التكرار	حدود الفئات	التكرار المتجمع الصاعد
1	—	—	أقل من 40	صفر
2	- 40	2	أقل من 46	2
3	- 46	10	أقل من 52	12
4	- 52	12	أقل من 58	24
5	- 58	15	أقل من 62	39
فئة الوسيط 				
6	- 62	20	أقل من 68	59
7	- 68	17	أقل من 76	76
8	- 76	11	أقل من 82	87
9	- 82	8	أقل من 88	95
10	- 88	4	أقل من 94	99
11	100 - 94	1	أقل من 100	100

أو باستخدام المعادلة (2-15) :

$$M = L + \frac{\frac{n}{2} - k_1}{F} \times i = 58 + 4 \left(\frac{50 - 39}{20} \right) = 58 + \frac{(4)(11)}{(20)}$$

$$\therefore M = 58 + 2.2 = 60.2$$

وهي نفس القيمة السابقة. ولتحقيق الدقة في حساب الوسيط ينصح باستخدام المعادلة (2-14) لحسابه .

مثال (2-18)

أوجد الوسيط لعمر مائة عضو في جمعية نسائية حسب البيانات المبينة في الجدول (2-21) :

جدول (2-21)

الفئة	- 20	- 25	- 30	- 35	- 40	- 45	- 50	55 - فما فوق
التكرار	3	9	13	16	20	15	13	11

الحل :

أن مجموع التكرارات = 100 إذن رتبة الوسيط تساوي :

$$50 = \frac{100}{2}$$

ثم نكمل جدول التكرار المتجمع الصاعد لهذه البيانات كما يبين الجدول (2-22) ، من جدول التكرار المتجمع الصاعد السابق حيث نلاحظ أن :

$$L_1 = 35 , L_2 = 40 , i = 40 - 35 = 5 , k_1 = 41 , F = 20 , k_2 = 61$$

وباستخدام المعادلة (2-14) أو المعادلة (2-15) نوجد قيمة الوسيط :

$$M = L + \frac{\frac{n}{2} - k_1}{k_2 - k_1} i = 35 + 5 \left(\frac{50 - 41}{61 - 41} \right) = 35 + 5 \left(\frac{9}{20} \right)$$

$$\therefore M = 35 + \frac{9}{4} = 35 + 2.25 = 37.25$$

جدول (22-2)

ف	فئات العمر	التكرار	حدود الفئات	التكرار المتجمع الصاعد
—	—	—	أقل من 20	صفر
1	- 20	3	أقل من 25	3
2	- 25	9	أقل من 30	12
3	- 30	13	أقل من 35	25
4	- 35	16	أقل من 40	41
<div style="display: flex; align-items: center; justify-content: center;"> <div style="width: 100px; border-bottom: 1px solid black; margin-bottom: 5px;"></div> <div style="margin-left: 10px;">→</div> </div> <p style="text-align: center;">فئة الوسيط</p>				
5	- 40	20	أقل من 45	61
6	- 45	15	أقل من 50	76
7	- 50	13	أقل من 55	89
8	55 فما فوق	11	أقل من 60	100
	Total المجموع	100	—	—

أما باستخدام المعادلة (15-2) :

$$M = L + \frac{\frac{n}{2} - k_1}{F} x i = 35 + 5 \left(\frac{50 - 41}{20} \right) = 35 + \frac{(5)(9)}{(20)}$$

$$\therefore M = 35 + 2.25 = 37.25$$

10.2 الربع الأدنى والربع الأعلى

(Quartile Lower Quartile & Upper)

هناك قيم أخرى سببها بالوسيط في طريقة حسابها ولكنها ليست من المتوسطات مثل الربع والعشيرة والثنين وغيرها . وسنكتفي هنا بدراسة الربع الأدنى (Lower Quartile) والربع الأعلى (Upper Quartile) نظراً لحاجتنا إليهما فيما بعد لقياس تشتت البيانات .

أن الربع الأدنى هو القيمة التي تقسم المجموعة إلى قسمين بحيث تقع 25% من القيم قبلها و 75% من القيم بعدها ، بشرط ترتيب القيم تصاعدياً ، والربع الأعلى هو القيمة التي تقسم مجموعة القيم إلى قسمين بحيث يقع 75% من القيم قبلها و 25% من القيم بعدها . ويمكن حساب الربعين من التوزيعات التكرارية بإتباع نفس الطريقة التي حسب بها الوسيط ففي حالة حساب الربع الأدنى لتوزيع درجات الذكاء للخمسين مهندس نتبع ما يلي :

1- نجد أولاً ترتيب الربع الأدنى والذي يساوي :

$$12.5 = \frac{50}{4} = \frac{n}{4}$$

2- نعين فئة الربع الأدنى من جدول التكرار المتجمع الصاعد للبيانات . ففي حالتنا هذه نجدها الفئة (100 وأقل من 110) .

3- نوجد قيمة الربع الأدنى باستخدام العلاقة التالية :

$$R_1 = L + \frac{\frac{n}{4} - k_1}{k_2 - k_1} i \dots\dots\dots (16-2)$$

حيث أن :

R_1 - الربيع الأدنى للبيانات .

L - الحد الأدنى لفئة الربيع الأدنى .

n - عدد القيم أو البيانات .

k_1 - التكرار المتجمع الصاعد السابق لترتيب الربيع الأدنى .

k_2 - التكرار المتجمع الصاعد اللاحق لترتيب الربيع الأدنى .

i - طول فئة الربيع الأدنى .

ولغرض تطبيق المعادلة (2-16) لإيجاد الربيع الأدنى لدرجات اختبار الذكاء للخمسين مهندس وعند مراجعة جدول التكرار المتجمع الصاعد ، نجد أن :

$$i = 110 - 100 = 10 \text{ أما } k_2 = 17 \text{ و } k_1 = 3$$

وعند تعويض القيم في المعادلة (2 - 16) نحصل على :

$$R_1 = L + \frac{\frac{n}{4} - k_1}{k_2 - k_1} i = 100 + 10 \left(\frac{12.5 - 3}{17 - 3} \right) = 100 + \frac{(10)(9.5)}{14}$$
$$\therefore R_1 = 100 + 6.8 \approx 106.8$$

4- وبنفس الطريقة يمكن إيجاد الربيع الأعلى لتلك البيانات من المعادلة (2-17) حيث :

$$R_3 = L + \frac{\frac{3n}{4} - k_1}{k_2 - k_1} i \dots\dots\dots (2-17)$$

حيث ان :

R_3 - الربع الأعلى للبيانات .

L - الحد الأدنى لفئة الربع الأعلى .

n - عدد القيم أو البيانات .

k_1 - التكرار المتجمع الصاعد السابق لترتيب الربع الأعلى .

k_2 - التكرار المتجمع الصاعد اللاحق لترتيب الربع الأعلى .

i - طول فئة الربع الأعلى .

لغرض تطبيق المعادلة (2-17) لإيجاد الربع الأعلى لدرجات اختبار الذكاء للخمسين مهندس وعند مراجعة جدول التكرار المتجمع الصاعد ، سنجد بأن فئة الربع الأعلى هي :

$$k_2 = 44 \text{ و } k_1 = 33 \text{ و } (130 - 120)$$

أما طول الفئة فسوف نجده يساوي :

$$i = 110 - 100 = 10$$

أما ترتيب الربع الأعلى فهو:

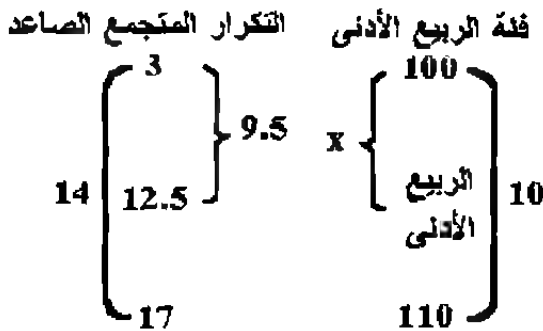
$$37.5 = \frac{150}{4} = \frac{3n}{4}$$

وعند تعويض القيم في المعادلة (2-17) نحصل على :

$$R_3 = L + \frac{\frac{3n}{4} - k_1}{k_2 - k_1} i = 120 + 10 \left(\frac{37.5 - 33}{44 - 33} \right) = 100 + \frac{(10)(4.5)}{11}$$

$$\therefore R_3 = 120 + 4.1 \approx 124.1$$

وبطريقة أخرى :



1- نجد ترتيب الربع الأدنى والذي يساوي $12.5 = \frac{50}{4}$.

2- من جدول التكرار المتجمع الصاعد نجد أن فئة الربع الأدنى هي الفئة (100 وأقل من 110) وطولها 10 وتكرارها 14 .

3- وبحسابنا لبعد ترتيب الربع الأدنى (12.5) عن التكرار المتجمع الصاعد السابق وهو 3 ، نجد أن:

$$9.5 = 3 - 12.5$$

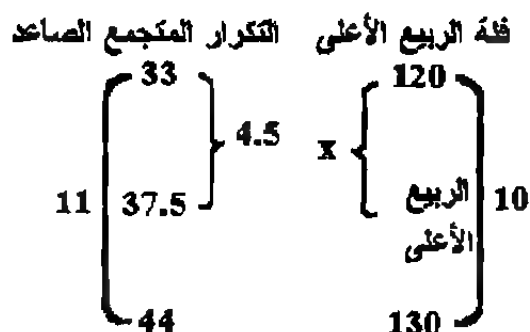
وهذا البعد يناظر بعد الربع الأدنى عن الحد الأدنى لفئة الربع الأدنى والذي سنرمز له بالرمز x حيث الربع الأدنى = $x + 100$.

4- إن البعد x يمكن الحصول عليه من العلاقة التالية :

$$\frac{x}{10} = \frac{9.5}{14} \Rightarrow \therefore 14x = 95 \Rightarrow \therefore x = \frac{95}{14} = 6.8$$

وعليه فإن الربع الأدنى يساوي $106.8 = 6.8 + 100$.

وبنفس الطريقة يمكن حساب الربع الأعلى كما موضح في الشكل :



1- نجد ترتيب الربع الأعلى والذي يساوي $37.5 = \frac{(50)(3)}{4}$.

2- من جدول التكرار المتجمع الصاعد نجد أن فئة الربع الأعلى هي الفئة (120 وأقل من 130) وطولها 10 وتكرارها 11 .

3- وبحسابنا لبعد ترتيب الربع الأعلى (37.5) عن التكرار المتجمع الصاعد السابق وهو 33 ، نجده يساوي :

$$4.5 = 33 - 37.5$$

وهذا البعد ينظر بعد الربع الأعلى عن الحد الأدنى لفئة الربع الأعلى والذي سنرمز له بالرمز x ، حيث أن الربع الأعلى = $x + 120$. إن البعد x نحصل عليه من العلاقة التالية :

$$\frac{x}{10} = \frac{4.5}{11} \Rightarrow \therefore 11x = 45 \Rightarrow \therefore x = \frac{45}{11} = 4.1$$

إن الربع الأعلى = $120 + 4.1 = 124.1$. ويجب الإشارة إلى أنه في كثير من الأحيان يطلق على الوسيط أيضاً بالربع الثاني R_2 .

مثال (2-19)

أوجد كل من الوسيط والربيعين الأدنى والأعلى للبيانات كما هو مبين في الجدول (2-23) :

جدول (2-23)

الترتيب	15 -	20 -	25 -	30 -	35 -	40 -	45 -	50 - 55
التكرار	3	4	9	8	11	15	2	6

الحل :

نكون أولاً جدول التكرار المتجمع الصاعد للبيانات والمبين في الجدول (2-24) .

جدول (2-24)

ت	الترتيب	التكرار	حدود الفئات	التكرار المتجمع الصاعد
—	—	—	أقل من 15	صفر
1	15 -	3	أقل من 20	3
2	20 -	4	أقل من 25	7
3	25 -	9	أقل من 30	16
4	30 -	8	أقل من 35	24
الترتيب الوسيطية				
5	35 -	11	أقل من 40	35
6	40 -	15	أقل من 45	50
7	45 -	2	أقل من 50	52
8	55 - 50	6	أقل من 55	58
المجموع		58	—	—

(a) لحساب الوسيط نقوم أولاً بإيجاد رتبة الوسيط والتي تساوي $29 = \frac{58}{2}$.
من جدول التكرار المتجمع الصاعد نجد أن :

$$L = 30, k_1 = 24, k_2 = 35, i = 5$$

وباستخدام المعادلة (2 - 14) نوجد الوسيط :

$$M = L + \frac{\frac{n}{2} - k_1}{k_2 - k_1} i = 30 + 5 \left(\frac{29 - 24}{35 - 24} \right) = 30 + 5 \left(\frac{5}{11} \right)$$

$$\therefore M = 30 + \frac{25}{11} = 30 + 2.27 = 32.27$$

(b) أم لحساب الربيع الأدنى نجد رتبة الربيع الأدنى حيث $14.5 = \frac{58}{4}$.
ومن جدول التكرار المتجمع الصاعد نجد أن :

$$L = 20, k_1 = 7, k_2 = 16, i = 5$$

بإستخدام المعادلة (2-16) نوجد الربيع الأدنى :

$$R_1 = L + \frac{\frac{n}{4} - k_1}{k_2 - k_1} i = 20 + 5 \left(\frac{14.5 - 7}{16 - 7} \right) = 20 + \frac{(5)(7.5)}{9}$$

$$\therefore R_1 = 20 + 4.167 \approx 24.167$$

(c) أما لحساب الربيع الأعلى نجد رتبة الربيع الأعلى والتي تساوي
 $43.5 = \frac{(58)(3)}{4}$.

من جدول التكرار المتجمع الصاعد نجد أن :

$$L = 35, k_1 = 35, k_2 = 50, i = 5$$

وباستخدام المعادلة (17-2) نجد الربع الأعلى :

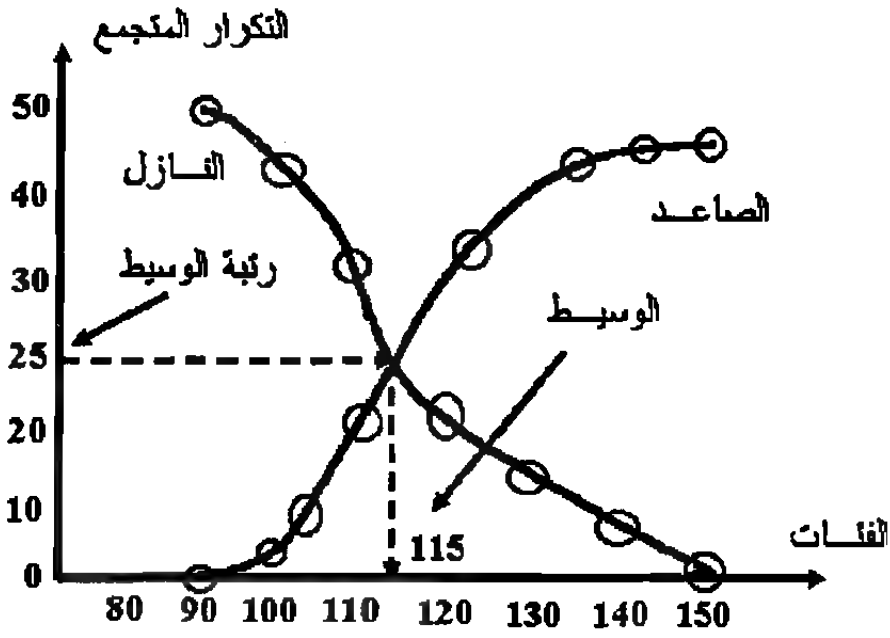
$$R_3 = L + \frac{\frac{3n}{4} - k_1}{k_2 - k_1} i = 35 + 5 \left(\frac{43.5 - 35}{50 - 35} \right) = 35 + \frac{(5)(8.5)}{15}$$
$$\therefore R_3 = 35 + \frac{8.5}{3} \approx 35 + 2.83 \approx 37.83$$

11.2 إيجاد الوسيط والربيعين بيانياً

(Finding Median and Quartiles Graphically)

لإيجاد الوسيط بيانياً نتبع الخطوات التالية :

- 1- نرسم منحني التكرار المتجمع الصاعد (أو النازل) للبيانات .
- 2- نعين ترتيب الوسيط على المحور الرأسي .
- 3- نرسم خطاً أفقياً حتى يقابل المنحني المتجمع في نقطة ما .
- 4- نسقط من تلك النقطة عموداً على المحور الأفقي فيقابله عند قيمة الوسيط .
- 5- وإذا رسمنا كلاً من منحني التكرار المتجمع الصاعد ومنحني التكرار المتجمع النازل في شكل واحد نجد أن المنحنيين يلتقيان في نقطة واحدة تحدد لنا قيمة الوسيط . إذا أسقطنا منها عموداً على المحور الأفقي فسوف نجد الفئة التي يقع فيها الوسيط ، والشكل (2-1) يوضح كيفية إيجاد الوسيط بالطريقة البيانية من تقاطع منحني التكرار المتجمع الصاعد ومنحني التكرار المتجمع النازل لدرجات اختبار الذكاء للخمسين مهندساً ونجد أن قيمة الوسيط يساوي 115 .



الشكل (1 - 2)

إيجاد الوسيط بيانياً لدرجات اختبار الذكاء للخمسين مهندس

ولإيجاد الربعين بيانياً نتبع نفس الخطوات التي أتبعنا لإيجاد الوسيط فإذا استخدمنا منحنى التكرار المتجمع الصاعد ، يكون ترتيب الربع الأدنى :

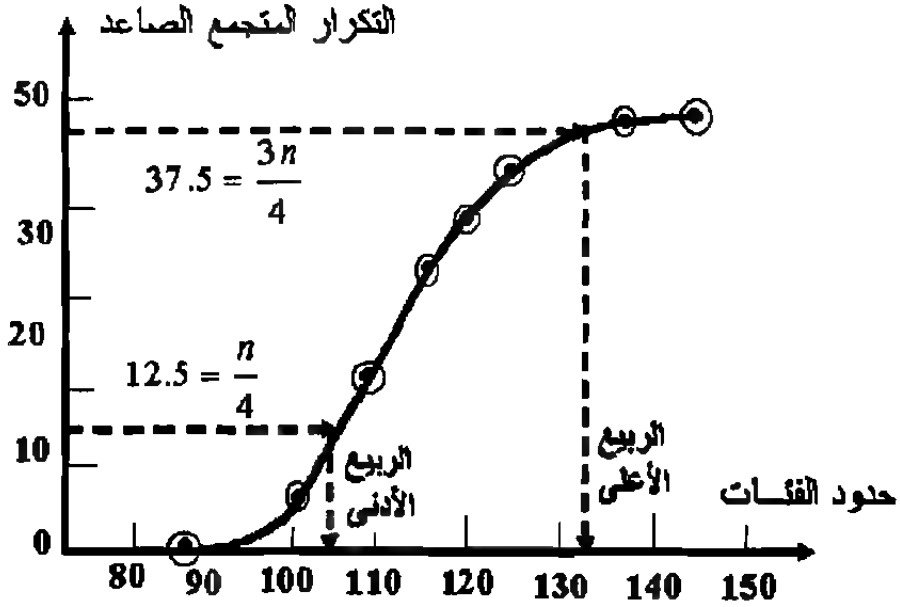
$$\frac{\sum f_i}{4}$$

ويكون ترتيب الربع الأعلى :

$$\frac{3(\sum f_i)}{4}$$

1- لإيجاد الربع الأدنى نعين ترتيبه على المحور الرأسي ونرسم خطاً أفقياً ليقابل المنحنى المتجمع الصاعد في نقطة. نسقط منها عموداً ليقابل المحور الأفقي عند قيمة الربع الأدنى ونجده في حالة درجات الذكاء للخمسين مهندس يساوي 107 درجة تقريباً كما موضح في الشكل (2-2) .

2- ولإيجاد قيمة الربع الأعلى نعين ترتيبه على المحور الرأسي ونتبع نفس الطريقة لتحديد قيمته على المحور الأفقي فنجد قيمته 124 درجة تقريباً لدرجات اختبار الذكاء للخمسين مهندس كما موضح في الشكل (2-2) .



الشكل (2-2)

إيجاد الربعين الأدنى والأعلى بيانياً لدرجات اختبار الذكاء للخمسين مهندس

12.2 خواص الوسيط (The Median Properties)

- 1- قيمة الوسيط على العكس من المتوسط الحسابي لا تتأثر بوجود القيم الشاذة أي القيم الكبيرة جداً أو الصغيرة جداً ، ففي حالة وجود مثل هذه القيم يفضل استخدام الوسيط .
- 2- يمكن حساب الوسيط من الجداول المفتوحة لأننا في حسابه لا نحتاج لمعرفة مراكز الفئات مثل المتوسط الحسابي .
- 3- يمكن إيجاد الوسيط باستخدام الرسم البياني .

إن الوسيط هو مؤشر مفيد في بعض الأحوال لمعرفة النزعة المركزية لمجموعة من البيانات مثل التوزيعات المركزية ذات الفترات المفتوحة والتوزيعات ذات القيم المتطرفة وطرق حسابه أبسط من طرق حساب المتوسط الحسابي ولكنه كما ذكرنا أقل ثباتاً من المتوسط من عينة لأخرى لذلك فإن استعماله محدودة مقارنة مع الاستخدامات الواسعة للمتوسط الحسابي . وكما أن المتوسط مؤشر أبسط فهناك مؤشر ثالث أبسط من الوسيط يعطينا فكرة سريعة عن النزعة المركزية لمجموعة من البيانات ذلك هو المنوال Mode .

13.2 المنوال (The Mode)

إن المنوال لمجموعة من البيانات هو القيمة التي تحظى بأكبر تكرار من هذه المجموعة من البيانات أو هو القيمة التي تحدث بصورة أكثر من غيرها وهو القيمة الأكثر شيوعاً فمثلاً إذا كان لدينا مجموعة القيم التالية :

1 , 5 , 3 , 12 , 7 , 4 , 3 , 9 , 12 , 4 , 3

نجد أن القيمة 3 تكررت أكثر من غيرها ولذلك نعتبر منوال المجموعة أما القيم التالية :

1 , 5 , 12 , 7 , 9 , 12 , 4 , 3

فلا يوجد لها منوال في حين أن القيم التالية :

1 , 5 , 12 , 7 , 5 , 3 , 9 , 4 , 3

نجد لها منوالين حيث إن كل من القيمتين 3 , 5 قد تكررت نفس العدد من المرات وقد تختلف قيمة المنوال المحسوبة من التوزيع التكراري عن قيمة المنوال للبيانات غير المبوبة كما أن اختلاف طول الفئات في التوزيعات التكرارية يؤدي أحياناً إلى تغيير موضع المنوال .

1.13.2 الخواص العامة للمنوال (General Properties for Mode)

هناك عدة خواص للمنوال من أهمها :

- 1- لا يتأثر بالقيم المتطرفة (الشاذة) التي قد توجد بين قيم المجموعة.
- 2- يمكن حسابه من الجداول التكرارية المفتوحة.
- 3- ليس له معنى إذا كانت التكرارات قليلة.
- 4- يعتبر من أحسن المتوسطات في وصف البيانات النوعية.
- 5- يمكن إيجاده من الرسم البياني كما سيتم توضيحه لاحقاً.
- 6- يتأثر المنوال بتغيير أطوال الفئات مما يقلل من أهميته ومن استخدامه وفي حالة المنحنيات التكرارية التي لها نهاية صغرى والمنحنيات ذات الفرع الواحد يصبح المنوال قيمة طرفية لا معنى لها.
- 7- يعتبر المنوال المقياس الوحيد الذي يمكن استخدامه لإيجاد متوسط الظواهر التي لا يمكن قياسها كمياً (مثل الصفات) حيث يمكن اعتبار الصفة الأكثر شيوعاً هي منوال المجموعة.

إن المنوال رغم استخدامه في بعض الأحيان كمؤشر سريع لمجموعة من البيانات أو لمعرفة القيمة الأكثر شيوعاً لأغراض اجتماعية أو اقتصادية إلا أنه أقل المتوسطات استخداماً في العمليات الإحصائية الأكثر عمقاً وذلك لأنه لا يأخذ بنظر الاعتبار كل القيم كما أنه يتغير من عينة إلى أخرى لنفس المجتمع .

2.13.2 طرق حساب المنوال

(Methods for Calculating The Mode)

أولاً : في حالة البيانات المفردة (For Single Data)

كما قلنا سابقاً فإن قيمة المنوال تحدد من القيمة الأكثر شيوعاً أو الأكثر تكراراً في البيانات والمثال (20-2) سوف يوضح لنا ذلك .

مثال (20-2)

أوجد المنوال للقيم التالية :

31, 36 , 31, 19 , 30 , 31, 35 , 36 , 31, 50 , 31

الحل :

المنوال هو القيمة الأكثر تكراراً وهو في هذه الحالة " $M_o = 31$ " .

ثانياً : في حالة البيانات المبوبة (In Case of Tabulated Data)

لإيجاد المنوال من الجداول التكرارية نتبع ما يلي :

(a) نحدد الفئة المنوالية وهي الفئة التي تقابل أكبر تكرار .

(b) نحدد التكرار ما قبل وبعد الفترة المنوالية .

(c) نطبق القانون التالي :

$$M_o = L + \frac{F_1}{F_2 + F_1} i \quad \dots\dots\dots (19-2)$$

حيث أن :

M_0 - المنوال .

F_1 - التكرار ما قبل الفئة المنوالية .

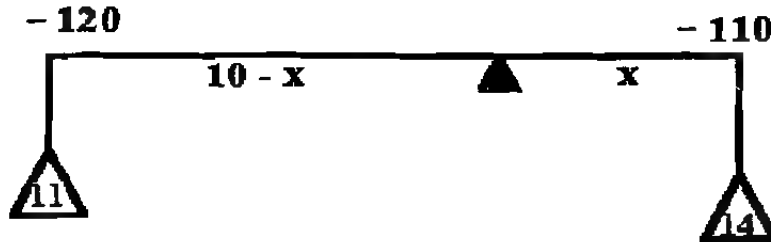
F_2 - التكرار ما بعد الفئة المنوالية .

L - الحد الأدنى للفئة المنوالية و i هو طول الفئة المنوالية .

أي نلاحظ الفئة التي يقع فيها المنوال وهي الفئة التي تحتوي على أكبر تكرار وتسمى بالفئة المنوالية (Mode Interval) ويقع المنوال عند مركز الفئة هذه إذا كان تكرار الفئة قبل المنوالية يساوي تكرار الفئة بعد المنوالية. ولتحديد موضع المنوال داخل الفئة المنوالية هناك طرق مختلفة لتحقيق ذلك منها :

1- طريقة الرافعة (The Crane Method)

لتحديد موضع المنوال داخل الفئة المنوالية ، تصوّر أن الفئة المنوالية عبارة عن رافعة (Crane) ، ويمثل تكرار الفئة قبل المنوالية القوة وتكرار الفئة بعد المنوالية المقاومة وعلى هذا الأساس يتحدد موضع المنوال عند موضع ارتكاز هذه الرافعة كما موضح في الشكل (2 - 3) .



الشكل (2- 3)

طريقة الرافعة لإيجاد المنوال لدرجات اختبار الذكاء للخمسين مهندس

أما لحساب المنوال للتوزيع التكراري الخاص باختبار الذكاء للخمسين مهندس نجد أن أكبر تكرار يقع أمام الفئة (110 وأقل من 120) أي أنها الفئة المنوالية . ولتحديد موضع المنوال داخلها ، نفرض أنه يبعد مسافة قدرها x عن بداية الفئة وهي القيمة 110 وبالتالي سيبعد مسافة قدرها (طول الفئة - x) أي $(10 - x)$ عن نهاية الفئة 120 وبالتالي نجد أن :

القوة x نراعاها = المقاومة x نراعاها

$$14x = 11(10 - x)$$

$$14x = 110 - 11x$$

$$\therefore 25x = 110 \Rightarrow \therefore x = \frac{110}{25} = 4.4$$

أي أن المنوال يساوي :

$$110 + 4.4 = 114.4 \text{ درجة.}$$

2- طريقة الفروق أو ما يعرف بطريقة العالم بيرسون

(The Differences or Person's Method)

أقترح " كارل بيرسون " أن الذي يحدد موضع المنوال داخل الفئة المنوالية هو الفرق بين تكرار الفئة المنوالية ، وتكراري الفئتين السابقتين واللاحقة لها . وعليه يتحدد موضع المنوال بحيث يقسم الفئة المنوالية إلى قسمين هما Δ_1 ; Δ_2 حيث أن Δ_1 هو الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها ، Δ_2 هو الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها . ولحل المثال الخاص باختبار درجات الذكاء للخمسين مهندس وإيجاد المنوال بطريقة بيرسون أي طريقة الفروق نجد أن :

$$5 - 11 - 16 = \Delta_2 \quad , \quad 2 - 14 - 16 = \Delta_1$$

والمنوال يكون القيمة التي تقسم الفئة المنوالية (110 وأقل من 120) بنسبة 5:2 فإذا كان بعد المنوال من بداية الفئة هو x فيكون بعده عن نهايتها $(x - 10)$ ، وهنا نجد أن $x : x - 10$ يجب أن تكون كنسبة 2 : 5 أي أن :

$$\frac{x}{10-x} = \frac{2}{5} \Rightarrow \therefore 5x = 2(10-x)$$

$$5x = 20 - 2x \Rightarrow \therefore 7x = 20 \Rightarrow \therefore x = \frac{20}{7} = 2.86$$

أي أن المنوال يساوي : $112.86 = 2.86 + 110$.

ويمكن الحصول على هذه القيمة مباشرة من العلاقة (20-2) حيث .

$$M_o = L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) i \dots \dots \dots (20-2)$$

حيث أن :

M_o و L و i قد تم تعريفهما مسبقاً في المعادلة (2 - 19) .

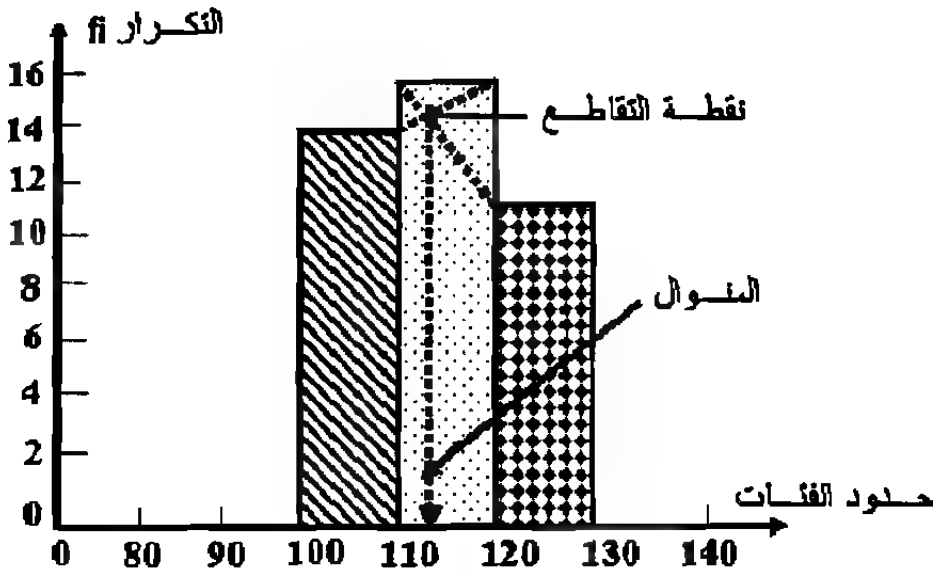
وبالتعويض نحصل على :

$$\begin{aligned} M_o &= L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) i = 110 + (10) \left(\frac{2}{5+2} \right) = 110 + \frac{20}{7} \\ &= 110 + 2.86 = 112.86 \end{aligned}$$

3- الطريقة البيانية (Graphical Method)

يمكن الحصول على المنوال بيانياً بأن نرسم من المدرج التكراري (Histogram)، الثلاثة مستطيلات التي تمثل تكرارات الفئة المنوالية والفئتين السابقتين واللاحقة كما موضح في الشكل (2-4)، ثم نصل الركن الأيمن العلوي للمستطيل الذي يمثل تكرار الفئة المنوالية بالركن الذي يماثله في المستطيل الذي يمثل تكرار الفئة التي قبلها.

ثم نصل الركن الأيسر العلوي للمستطيل الذي يمثل تكرار الفئة المنوالية بالركن الذي يماثله في المستطيل الذي يمثل تكرار الفئة بعد المنوالية، فيتقابل المستقيمان في نقطة. نسقط منها عموداً على المحور الأفقي فيحدد لنا موقع المنوال.



الشكل (2-4)

إيجاد المنوال بيانياً لاختبار الذكاء للخمسين مهندس

أحسب المنوال بالطريقة الجبرية والبيانية وطريقة بيرسون للجدول التكراري (24-2) :

جدول (24-2)

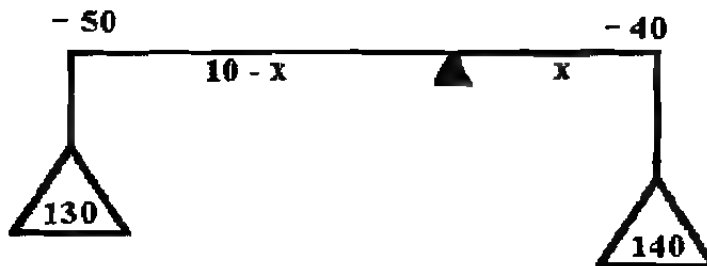
100 - 90	- 80	- 70	- 60	- 50	- 40	- 30	- 20	- 10	- 1	تكرار
30	40	70	100	130	170	140	80	40	10	التكرار

الحل :

إن الفئة المنوالية هي التي تقابل أكبر تكرار وهي (50 - 40) .

(1) طريقة الرافعة أو الطريقة الجبرية :

الحد الأدنى للفئة المنوالية هو $L = 40$ ، طول الفئة المنوالية = 10 تكرار الفئة المنوالية = 170 ، وتكرار الفئة قبل المنوالية يساوي 140 ، وتكرار الفئة بعد المنوالية = 130 ، نفرض بعد المنوال عن الحد الأدنى للفئة المنوالية = x ، إذن بعد المنوال عن الحد الأعلى للفئة المنوالية يساوي $x - 10$. وباستخدام طريقة الرافعة كما موضح في الشكل الآتي نجد أن :



$$140 x = 130 (10 - x)$$

$$140 x = 1300 - 130 x$$

$$\therefore 270 x = 1300 \Rightarrow \therefore x = \frac{1300}{270} \approx 4.82$$

أي أن المنوال يساوي :

$$44.82 = 4.82 + 40$$

(2) طريقة الفروق (طريقة العالم بيرسون) :

$$40 = 130 - 170 = \Delta_2 \quad , \quad 30 = 140 - 170 = \Delta_1$$

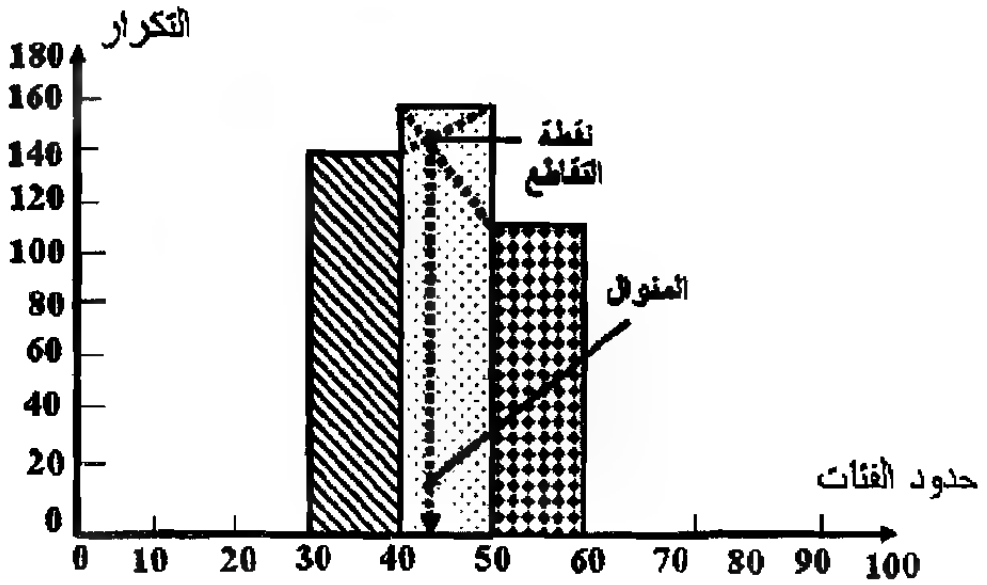
ونعوض في المعادلة رقم (2-20) فنحصل على قيمة المنوال حيث :

$$\begin{aligned} M_o &= L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) i = 40 + (10) \left(\frac{30}{30 + 40} \right) = 40 + \frac{300}{70} \\ &= 40 + 4.286 = 44.286 \end{aligned}$$

(3) الطريقة البيانية :

كما أشرنا في البند السابق ، نرسم من المدرج التكراري الثلاثة مستطيلات التي تمثل تكرارات الفئة المنوالية والفئتين السابفة واللاحقة ، ثم نصل الركن الأيمن العلوي للمستطيل الذي يمثل تكرار الفئة المنوالية بالركن الذي يمثله في المستطيل الذي يمثل تكرار الفئة التي قبلها . ثم نصل الركن الأيسر العلوي للمستطيل الذي يمثل تكرار الفئة المنوالية بالركن الذي يمثله في المستطيل الذي يمثل تكرار الفئة بعد المنوالية ، فيتقابل المستقيمان في

نقطة . نسقط منها عموداً على المحور الأفقي فيحدد لنا موقع المنوال ومن الرسم في الشكل (5-2) ، يتبين أن مسقط نقطة تقاطع المستقيمين على المحور الأفقي هي 44.2 تقريباً والتي تمثل قيمة المنوال وهي قيمة معاربة للقيمتين المحسوبتين في الطريقتين السابقتين والشكل (2 - 5) يوضح ذلك .



الشكل (2 - 5)

إيجاد المنوال بيانياً

4- إيجاد المنوال بيانياً للتوزيعات التكرارية غير المنتظمة :

عند إيجاد المنوال للبيانات المبوبة في جداول تكرارية غير منتظمة أي في جداول تكرارية أطوال فئاتها غير متساوي ، يجب إجراء عملية تعديل التكرارات التي سبق شرحها عند دراستنا للمدرج التكراري قبل إيجاد المنوال بأية طريقة من الطرق السابقة .

14.2 الوسط الهندسي والوسط التوافقي والعشيرات والمئينات (Geometric Mean, Harmonic Mean, Deciles and Percentiles)

1.14.2 الوسط الهندسي (G) - (Geometric Mean)

الوسط الهندسي لمجموعة من القيم الموجبة عددها n يساوي الجذر النوني لحاصل ضرب هذه القيم . فإذا كانت هذه القيم هي $x_1, x_2, x_3, \dots, x_n$ فإن الوسط الهندسي لهذه القيم G يحسب من المعادلة التالية :

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdots x_n} \quad \dots\dots\dots(21-2)$$

وإذا قمنا بأخذ لوغاريتم العلاقة (21-2) فالتنا نحصل على :

$$\begin{aligned} \log G &= \log [(x_1 \cdot x_2 \cdot x_3 \cdots x_n)^{\frac{1}{n}}] \\ \therefore \log G &= \frac{1}{n} \log [x_1 \cdot x_2 \cdot x_3 \cdots x_n] = \frac{1}{n} [\log x_1 + \log x_2 + \cdots + \log x_n] \\ \therefore \log G &= \frac{1}{n} \sum_{i=1}^n \log x_i \quad \dots\dots\dots(22-2) \end{aligned}$$

أي أن لوغاريتم الوسط الهندسي لمجموعة من القيم يساوي الوسط الحسابي للوغاريتمات هذه القيم .

فإذا كانت لدينا مثلاً القيم التالية :

122, 116, 126, 130, 116, 120, 132, 107

فيمكن عندئذٍ حساب الوسط الهندسي بالشكل التالي :

$$\begin{aligned} \log G &= \frac{1}{8} [\log x_1 + \log x_2 + \log x_3 + \log x_4 + \log x_5 + \dots + \log x_n] \\ \log G &= \frac{1}{8} [\log 107 + \log 132 + \log 120 + \log 116 + \\ &\quad \log 130 + \log 126 + \log 116 + \log 122] \\ &= \frac{1}{8} [2.0294 + 2.1206 + 2.0792 + 2(2.0645) \\ &\quad + 2.1139 + 2.1004 + 2.0864] \\ &= \frac{16.6589}{8} = 2.0824 \\ \therefore G &= 120.893 \end{aligned}$$

مثال (21-2)

أوجد الوسط الهندسي للسرعات الثلاثة التالية :

150 , 100 , 50

الحل :

$$G = \sqrt[3]{(50)(100)(150)} = \sqrt[3]{750,000} = 90.8$$

إيجاد الوسط الهندسي للبيانات المبوبة

(Geometric Mean for Data Set)

لإيجاد الوسط الهندسي G للبيانات المبوبة في شكل جداول تكرارية

نستخدم المعادلة التالية :

$$G = \sqrt[n]{\bar{x}_1^{f_1} \cdot \bar{x}_2^{f_2} \dots \bar{x}_n^{f_n}} \dots\dots\dots (23-2)$$

حيث أن \bar{x}_i هي مراكز الفئات و f_i هي تكرارات الفئات المقابلة للمراكز .

$$\begin{aligned}
\therefore \text{Log } G &= \text{Log} [\bar{x}_1^{f_1} \cdot \bar{x}_2^{f_2} \cdot \dots \cdot \bar{x}_n^{f_n}]^{\frac{1}{\sum f_n}} \\
&= \frac{1}{\sum f_n} [f_1 \text{Log } \bar{x}_1 + f_2 \text{Log } \bar{x}_2 + \dots + f_n \text{Log } \bar{x}_n] \\
\therefore \text{Log } G &= \frac{\sum f_n \text{Log } \bar{x}_n}{\sum f_n} \dots\dots\dots(24 - 2)
\end{aligned}$$

وتكون خطوات الحل كالآتي :

- 1- نعين مراكز الفئات .
- 2- نوجد قيم لوغاريتمات مراكز الفئات .
- 3- نضرب لوغاريتم مركز كل فئة في تكرار هذه الفئة .
- 4- نجمع حاصل الضرب فنحصل على :

$$\sum_{i=1}^n f_i \text{Log } \bar{x}_i$$

وبقسمتها على $\sum f_n$ نحصل على Log G .

5- باستخدام الحاسبة أو بيانات جداول اللوغاريتمات نجد قيمة الوسط الهندسي .

وكمثال على ذلك يمكن إيجاد الوسط الهندسي للجدول التكراري الخاص باختبار درجات الذكاء للخمسين مهندس كما يبين الجدول (25-2) حيث أن :

$$\text{Log } G = \frac{\sum_{i=1}^n f_i \text{Log } \bar{x}_i}{\sum_{i=1}^n f_i} = \frac{103.1110}{50} = 2.0622 \Rightarrow \therefore G = 115.4^\circ$$

جدول (2 - 25)

إيجاد الوسط الهندسي لدرجة المهندس في اختبار الذكاء

الفئات	التكرار f_i	مركز الفئة \bar{x}_i	لوغاريتم مركز الفئة $\text{Log } \bar{x}_i$	$f_i \text{ Log } \bar{x}_i$
90 -	3	95	1.9777	5.9331
-100	14	105	2.0212	28.2968
-110	16	115	2.0607	32.9712
-120	11	125	2.0969	23.0659
-130	4	135	2.1303	8.5214
150 - 140	2	145	2.1614	4.3228
المجموع	50	—	—	103.1110

والوسط الهندسي قليل الاستخدام ، إذا ما قورن بالمتوسط الحسابي ويستخدم عادة لإيجاد مجموعة من النسب وهو لذلك يستخدم في حساب الأرقام القياسية كما سنرى فيما بعد .

2.14.2 الوسط التوافقي (H) - (Harmonic Mean)

يعرف الوسط التوافقي بأنه عبارة عن مقلوب المتوسط الحسابي لمقلوب القيم ويفضل استخدام هذا الوسط في حساب معدل السرعة إذ أنها تعطى بدلالة الزمن .

طرق حساب الوسط التوافقي

أولاً في حالة البيانات المفردة يستخدم القانون التالي :

$$H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} \dots\dots\dots(25-2)$$

حيث أن n هو عدد القيم و x_i هي القيمة المفردة في مجموعة البيانات .

مثال (22-2)

أوجد الوسط التوافقي للقيم التالية: 6,2,10,4,8

الحل :

باستخدام المعادلة (22 - 2) نحصل على الوسط التوافقي كما يلي :

$$H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} = \frac{5}{\frac{1}{6} + \frac{1}{2} + \frac{1}{10} + \frac{1}{4} + \frac{1}{8}} = \frac{5}{1.1417} = 4.37$$

مثال (23 2)

أحسب الوسط التوافقي للقيم التالية: 10 - 22 - 8 - 15 - 18 - 13

الحل:

$$H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} = \frac{6}{\frac{1}{10} + \frac{1}{22} + \frac{1}{8} + \frac{1}{15} + \frac{1}{18} + \frac{1}{13}} = \frac{6}{0.461} = 15.01$$

ثانياً في حالة البيانات المبوبة تستخدم المعادلة التالية :

$$H = \frac{\sum_{i=1}^n f_i}{\sum_{i=1}^n \left(\frac{f_i}{\bar{x}_i} \right)} \dots\dots\dots (26 - 2)$$

حيث أن \bar{x}_i هي مراكز الفئات و f_i هو تكرار الفئة . وكمثال على ذلك يبين الجدول (26 - 2) كيفية إيجاد الوسط التوافقي لدرجات الذكاء للخمسين مهندس .

جدول (26-2)

إيجاد الوسط التوافقي لدرجات الذكاء للخمسين مهندس

الفئات	التكرار f_i	مركز الفئة \bar{x}_i	f_i / \bar{x}_i
90 -	3	95	0.0316
-100	14	105	0.1333
-110	16	115	0.1391
-120	11	125	0.088
-130	4	135	0.0296
150 - 140	2	145	0.0138
المجموع	50	—	0.4354

$$\therefore H = \frac{\sum_{i=1}^n f_i}{\sum_{i=1}^n \left(\frac{f_i}{\bar{x}_i} \right)} = \frac{50}{0.4354} = 114.837$$

خواص الوسط التوافقي (Harmonic Mean Properties)

هناك عدة خواص عامة للوسط التوافقي من أهمها :

- 1- يتأثر بالقيم المتطرفة حاله كحال المتوسط الحسابي .
- 2- لا يمكن حسابه من الجداول المفتوحة مثل المتوسط الحسابي .
- 3- قابل للعمليات الجبرية .
- 4- يستخدم في وصف تغيرات الظواهر النسبية .
- 5- دائماً يكون $H \leq G \leq \bar{x}$ ويحدث هذا التناوب عندما تكون جميع القيم متساوية .

مثال (24-2)

أوجد الوسط التوافقي للجدول التكراري (27-2) :

جدول (27-2)

27- 23	-19	- 15	- 11	-7	- 3	الفئات x_i
5	6	9	10	12	8	التكرار f_i

الحل :

نرتب الجدول (27-2) كما هو مبين في الجدول (28-2) ومنه نجد أن الوسط التوافقي يساوي :

جدول (28-2)

الفئات	التكرار f_i	مركز الفئة \bar{x}_i	f_i / \bar{x}_i
- 3	8	5	$1.6 = \frac{8}{5}$
- 7	12	9	$1.33 = \frac{12}{9}$
- 11	10	13	$0.77 = \frac{10}{13}$
- 15	9	17	$0.53 = \frac{9}{17}$
- 19	6	21	$0.286 = \frac{6}{21}$
27 - 23	5	25	$0.2 = \frac{5}{25}$
المجموع	50	—	4.716

$$H = \frac{\sum_{i=1}^n f_i}{\sum_{i=1}^n \left(\frac{f_i}{\bar{x}_i} \right)}$$

$$= \frac{50}{4.716} = 10.6$$

مثال (25-2)

أوجد الوسط التوافقي للتوزيع التكراري كما هو مبين في الجدول (29-2) :

جدول (29-2)

الفئات x_i	التكرار f_i
14- 12	9
-10	5
- 8	15
- 6	4
-4	1
- 2	3

الحل :

نرتب الجدول (29-2) كما هو مبين في الجدول (30-2) ومنه نستطيع أن نجد الوسط التوافقي والذي يساوي :

جدول (30-2)

الصفات	التكرار f_i	مركز الفئة \bar{x}_i	f_i / \bar{x}_i
- 2	3	3	$1.0 = \frac{3}{3}$
- 4	1	5	$0.2 = \frac{1}{5}$
- 6	4	7	$0.57 = \frac{4}{7}$
- 8	15	9	$1.67 = \frac{15}{9}$
- 10	5	11	$0.45 = \frac{5}{11}$
14 - 12	9	13	$0.69 = \frac{9}{13}$
المجموع	37	—	4.58

$$H = \frac{\sum_{i=1}^n f_i}{\sum_{i=1}^n \left(\frac{f_i}{\bar{x}_i} \right)}$$

$$= \frac{37}{4.58} = 8.1$$

3.14.2 العشريرات (Deciles)

ويرمز لها بالرمز Q_r وتحسب في حالة البيانات المبوبة من المعادلة التالية :

$$Q_r = L + \frac{\frac{n_r}{10} - k}{F} i \dots \dots \dots (27-2)$$

حيث أن :

- D_r - تساوي العشير .
- L - تساوي الحد الأدنى لفئة العشير .
- k - تساوي التكرار المتجمع الصاعد السابق لفئة العشير .
- i - تساوي طول فئة العشير .
- F - تساوي تكرار فئة العشير .

مثال (27-2)

أوجد D_1 , D_5 , D_8 للتوزيع التكراري المبين في الجدول (31-1) :

جدول (31-1)

20 - 18	- 16	- 14	- 12	- 10	- 8	- 6	- 4	- 2	الفئة
5	3	21	10	11	20	15	7	5	التكرار

الحل :

نبدأ بعمل جدول التكرار المتجمع الصاعد ومنه نجد العشير الأول D_1 والعشير الخامس D_5 والعشير الثامن D_8 كما هو مبين في الجدول (32-2) .

جدول (2-32)

	الحدود العليا للفئات	التكرار المتجمع الصاعد
	تقل من 2	0
	تقل من 4	5
D_1 - العنصر الأول	تقل من 6	12
	تقل من 8	27
	تقل من 10	47
D_5 - العنصر الخامس	تقل من 12	58
	تقل من 14	68
D_8 - العنصر الثامن	تقل من 16	89
	تقل من 18	92
	تقل من 20	97

أن رتبة العنصر الأول هي :

$$D_1 = \frac{(97)(1)}{10} = 9.7$$

وبما أن طول فئة العنصر يساوي 2 أي أن :

$$\therefore D_1 = L + \frac{\frac{n_r}{10} - k}{F} i = 4 + \frac{9.7 - 5}{7} \times 2$$

$$\therefore D_1 = 5.34$$

أما رتبة العنصر الخامس تساوي :

$$D_5 = \frac{(97)(5)}{10} = 48.5$$

لذلك فإن :

$$\therefore D_5 = L + \frac{\frac{n_r}{10} - k}{F} i = 10 + \frac{48.5 - 47}{11} \times 2$$

$$\therefore D_5 = 10.27$$

أن رتبة العشير الثامن هي :

$$D_8 = \frac{(97)(8)}{10} = 77.6$$

لذلك فإن :

$$\therefore D_8 = L + \frac{\frac{n_r}{10} - k}{F} i = 14 + \frac{77.6 - 68}{21} \times 2$$

$$\therefore D_8 = 14.9$$

4.14.2 المئينات (Percentiles)

ويرمز لها بالرمز M_r وتحسب في حالة البيانات المبوبة من المعادلة التالية :

$$M_r = L + \frac{\frac{n_r}{100} - k}{F} i \quad \dots\dots\dots(28-2)$$

حيث أن :

M_r - المئين .

L - الحد الأدنى لفئة المئين .

k - التكرار المتجمع الصاعد السابق لفئة المئين

i- طول فئة المئين .

F- تكرار فئة المئين .

مثال (28-2)

أوجد M_{20} , M_{60} , M_{90} للتوزيع التكراري كما هو مبين في الجدول (33-2) :

جدول (33-2)

الفئة	-125	-150	-175	-200	-225	275 - 250	المجموع
التكرار	16	164	400	320	93	7	1000

الحل :

نبدأ بعمل جدول التكرار المتجمع الصاعد ومنه نجد المئين العشرين M_{20} والمئين الستين M_{60} والمئين التسعين M_{90} كما مبين في الجدول (34-2) حيث أن رتبة المئين العشرين :

$$M_{20} = \frac{(1000)(20)}{100} = 200$$

وبما أن طول فئة المئين تساوي 25 فإن :

$$M_{20} = L + \frac{\frac{n_r}{100} - k}{F} i = 175 + \frac{200 - 180}{400} \times 25 = 176.25$$

جدول (2-34)

	الحدود العليا للفئات	التكرار المتجميع الصاعد
	أقل من 125	0
	أقل من 150	16
M_{20} - المئين العشرين	أقل من 175	180
M_{60} - المئين الستين	أقل من 200	480
M_{90} - المئين التسعين	أقل من 225	900
	أقل من 250	993
	أقل من 275	1000

أما رتبة المئين الستين فتساوي :

$$M_{60} = \frac{(1000)(60)}{100} = 600$$

وبما أن طول فئة المئين يساوي 25 فإن :

$$M_{60} = L + \frac{\frac{n_r}{100} - k}{F} i = 200 + \frac{600 - 480}{320} \times 25 = 209.37$$

أما رتبة المئين التسعين فتساوي :

$$M_{90} = \frac{(1000)(90)}{100} = 900$$

وبما أن طول فئة المئين يساوي 25 فإن :

$$M_{90} = L + \frac{\frac{n_r}{100} - k}{F} i = 200 + \frac{900 - 480}{320} \times 25 = 232.8$$

15.2 العلاقة بين المتوسطات (Relation between Means)

في حالة المنحنيات التكرارية المتماثلة نجد أن المتوسط الحسابي يساوي الوسيط ويساوي المنوال . وإذا كان المنحني التكراري قريباً جداً من التماثل نجد علاقة تقريبية تربط بين المتوسطات الثلاثة وهي :

$$\text{المتوسط الحسابي} - \text{المنوال} = 3 (\text{المتوسط الحسابي} - \text{الوسيط})$$

أو :

$$\text{Arithmetic Mean} - \text{Mode} = 3 (\text{Arithmetic Mean} - \text{Median})$$

ويلاحظ هنا أن الوسيط يقع بين المتوسط الحسابي والمنوال ، ويمكن استخدام هذه العلاقة في تقدير قيمة الوسط الحسابي من الجداول المفتوحة وذلك بواسطة حساب كل من الوسيط والمنوال .

مثال (2 - 29)

فصل دراسي فيه 42 طالب في مادة الإحصاء في أحد المراكز المهنية العليا. جلس منهم 40 في أحد امتحانات المادة وتغيب اثنان بسبب المرض فكان المتوسط الحسابي لدرجاتهم 67 من 100 . وبعد أسبوع تقدم الطالبان اللذان تغيبا عن الامتحان في المادة نفسها فتم إعادة الامتحان لهما فحصل

الأول على درجة 45 وحصل الثاني على 90 . فكم يصبح المتوسط الحسابي لدرجات جميع طلبة هذا الفصل الدراسي . وإذا قدرت قيمة المنوال لهذا الفصل 56 درجة كم هي قيمة الوسيط .

الحل :

بما أن المتوسط الحسابي يعرف على أنه هو مجموع القيم على عددها لذلك فإن 40 طالب الذين أدوا الامتحان في وقته قد جمعوا درجات مساوية إلى :

$$67 \times 40 = 2680$$

وعندما قام الطالبان اللذان تخلفا عن الامتحان في وقته بإعادة الامتحان وحصلوا على الدرجات المبينة في المسألة لذا فإن مجموع درجات الفصل يصبح كالتالي :

$$90 + 45 + 2680 = 2815$$

وعلى هذا الأساس يصبح المتوسط الحسابي كما يلي:

$$\bar{x} = \frac{2815}{42} = 67.024$$

ولغرض حساب قيمة الوسيط نطبق العلاقة التقريبية بين المتوسطات حيث :

$$\text{الوسط الحسابي} - \text{المنوال} = 3 (\text{الوسط الحسابي} - \text{الوسيط})$$

أي أن :

$$Mean - Mode = 3 (Mean - Median)$$

$$\therefore 67.024 - 30 = 3 (67.024 - M)$$

$$\therefore M = \frac{201.072 - 37.024}{3} = 54.683$$

16.2 تمارين

س1: يمثل الجدول التكراري (2-35) توزيع مائة عامل حسب الأجر الأسبوعي للعامل بالدنانير .

جدول (2-35)

الفئات	- 20	- 30	- 40	- 50	- 60	70 وأقل من 80
التكرار	10	15	30	22	14	9

أوجد ما يلي :

(a) الوسط الحسابي باستخدام طريقة الانحرافات المختصرة .

(b) الوسيط .

(c) المنوال بطريقة بيرسون .

(d) الوسط الهندسي .

س2: الجدول (2-36) يبين متوسط أوزان خمسة مجموعات من الطلبة حيث :

جدول (2-36)

عدد الطلبة	13	27	33	18	90
الوزن ، كجم	50	55	60	65	70

والمطلوب إيجاد الآتي :

(a) الوسط الحسابي المرجح للقيم .

(b) الوسط الهندسي لأوزان الطلبة في المجموعات الخمسة .

س3: أوجد المتوسط الحسابي والوسيط والمنوال والوسط الهندسي للتوزيع التكراري المبين في الجدول (37-2) :

جدول (37-2)

الفئات	- 12.5	- 22.5	- 32.5	- 42.5	52.5 وأقل من 62.5
التكرار	6	17	43	26	8

س4: أوجد كلاً من الوسط الحسابي والوسيط والمنوال للقيم التالية :

21 , 18 , 26 , 15 , 21 , 16 , 9 , 5 , 7 , 3

س5: التوزيع التكراري المبين في الجدول (38-2) يبين عدد مرات التوقف لآلة ما في مائة يوم عمل متتالية والمطلوب :

- أحسب متوسط مدة التوقف مستخدماً كل من المتوسط الحسابي والوسيط .
- هل من المفضل اختيار مقياساً آخرأً خلاف هذين المقياسين مع تفسير ذلك .
- أوجد الربيعين الأدنى والأعلى والمنوال للتوزيع التكراري .

جدول (38-2)

مدة توقف , دقيقة	- 0	- 10	- 20	- 30	- 40	- 50	- 60	- 70	80 - 90
عدد مرات التوقف	3	14	28	22	14	10	4	2	3

س6: الجدول (2-39) يوضح فئات أعمار 250 شخصاً يعملون في أحد الشركات الصناعية .

جدول (2-39)

فئة العمر، سنة	- 25	- 28	- 31	- 34	- 37	- 40	- 43	- 46	- 49
عدد الأشخاص	2	13	22	39	75	47	27	18	7

أوجد كل من المتوسط الحسابي والوسيط والمنوال مقرباً لأقرب سنة .

س7: فيما يلي التوزيع التكراري (2-40) للأجور اليومية لعدد من العاملين لمؤسسة ما :

جدول (2-40)

فئات الأجور بالدينار	- 60	- 120	- 180	- 240	- 300	المجموع
عدد العاملين	3	5	20	10	12	50

أحسب متوسط أجور العاملين مستخدماً طريقة الوسط الفرضي .

س8: جمعت بيانات عن الأجور الأسبوعية للعاملين في أربعة مصانع تعمل كلها في صناعة واحدة ووجد أن :

(a) عدد عمال المصنع الأول 300 والوسط الحسابي لأجورهم الأسبوعية 45 دينار .

(b) عدد عمال المصنع الثاني 600 والوسط الحسابي لأجورهم الأسبوعية 50 دينار .

(c) عدد عمال المصنع الثالث 450 والوسط الحسابي لأجورهم الأسبوعية 63 دينار .

(d) عدد عمال المصنع الرابع 300 موزعين حسب أجورهم الأسبوعية كما هو مبين في الجدول (2-41) :

جدول (2-41)

فئات الأجور بالدينار	-30	-40	-50	-60	-70	-80	90 وأقل من 100
عدد العمال	16	25	45	60	53	43	34

المطلوب ما يلي :

- (a) إيجاد المتوسط الحسابي للأجور اليومية لعمال المصنع الأربعة .
(b) إيجاد الوسيط و المنوال و الربيعين الأدنى والأعلى لعمال المصنع الرابع .

س9: أرسم المدرج التكراري للبيانات المبينة في الجدول (2-42) ومنه قدر قيمة المنوال :

جدول (2-42)

الفئات	-5	-10	-20	-50	-100	200 وأقل من 300
التكرار	26	94	369	380	410	180

ثم تحقق من القيمة التي يتم الحصول عليها من الرسم البياني حسابياً .

س10: الجدول (2-42) يبين العمر بالسنة الذي تزوج عنده مجموعة من الرجال .

جدول (2-42)

سن الزواج	- 15	- 20	- 25	- 30	- 35	- 40	- 45	50 فما فوق
التكرار	15	44	46	13	10	13	7	2

المطلوب أيجاد :

- (a) أستخدم أي مقاييس تراها مناسبة لإيجاد متوسط سن الزواج في هذه المجموعة ثم علل سبب اختيارك للمقاييس التي استخدمتها .
- (b) أحسب الربيع الأدنى والأعلى والعشير الأول والعشير السابع والمئين الخامس والثلاثين والمئين التاسع والسبعين في التوزيع المذكور .
- (c) أوجد كلاً من الوسيط والمنوال بالطريقة البيانية .

س11: الجدول (2-43) يبين التوزيع التكراري لخمسين طالباً حسب أوزانهم :

جدول (2-43)

فئات الوزن بالكيلوجرام	- 50	- 55	- 60	- 65	- 70	- 75	80 فأكثر
عدد الطلبة	3	5	11	16	9	4	2

المطلوب إيجاد :

- (a) أوجد كل من الوسيط والمنوال لجدول التوزيع التكراري .
- (b) أوجد الوسط الحسابي للتوزيع باستخدام العلاقة بين المتوسطات .

س12: أوجد المتوسط الحسابي والوسط الهندسي والوسط التوافقي للتوزيع كما هو مبين في الجدول (44-2) :

جدول (44-2)

112 وتكرار من 119	- 105	- 98	- 91	- 84	- 77	- 70	الترددات
4	11	17	30	21	11	6	التكرار

كذلك أوجد كلاً من المنوال والوسيط والربيع الأدنى والربيع الأعلى بيانياً وحسابياً للتوزيع التكراري .

س13: في جدول التوزيع التكراري (45-2) أوجد ما يلي :

جدول (45-2)

68-64	- 60	- 56	- 52	- 48	- 44	- 40	- 36	- 32	الترددات
5	41	26	35	45	25	14	12	9	التكرار

(a) الوسط الحسابي بطريقة الوسط الفرضي .

(b) الوسيط .

(c) المنوال .

(d) الربيع الأدنى .

(e) الربيع الأعلى .

(f) العشير الرابع والعشير السادس .

(g) المئين العشرين والمئين الثمانين .

(h) الوسط الهندسي والوسط التوافقي للقيم .

الباب الثالث

مقاييس التشتت (Measures of Dispersion)

- 1.3 مقدمة .
- 2.3 المدى .
- 3.3 الانحراف الربيعي .
- 4.3 الانحراف المتوسط .
- 1.4.3 الخواص العامة للانحراف المتوسط .
- 2.4.3 طرق حساب الانحراف المتوسط .
- 5.3 الانحراف المعياري .
- 6.3 إيجاد الانحراف المعياري للبيانات المبوبة .
- 7.3 إيجاد الانحراف المعياري بطريقة الانحرافات .
- 8.3 إيجاد الانحراف المعياري بطريقة الانحرافات المختصرة .
- 9.3 معامل الاختلاف .
- 10.3 الوحدات المعيارية .
- 11.3 المنحنى التكراري المتماثل والتوزيعات الاعتدالية .
- 12.3 الحيود عن التوزيع الاعتدالي .
- 13.3 تمارين .

يعرف التشتت (Dispersion) في أي مجموعة من القيم على أنه درجة التفاوت أو الاختلاف بين قيم هذه المجموعة ، فإذا كانت قيم المجموعة متقاربة من بعضها البعض يكون التشتت صغيراً وإذا كانت متباعدة عن بعضها البعض أي متباعدة يكون التشتت كبيراً .

ولمقارنة مجموعتين من البيانات لا نكتفي بمقاييس التوسط التي سبق لنا دراستها حيث قد يكون للمجموعتين نفس المتوسط الحسابي ولكنهما يختلفان في درجة التشتت ، فقد نجد أن مفردات إحدى المجموعتين متجمعة حول متوسط المجموعة بينما مفردات المجموعة الأخرى متناثرة ومتباعدة عن متوسطاتها ، وعندئذ يقال أن المجموعة الأولى أقل تشتتاً من المجموعة الثانية كما يتضح من المثال التالي :

$$x : 12 , 17 , 20 , 25 , 26$$

$$y : 8 , 12 , 20 , 26 , 34$$

حيث نجد أن المتغيرين x , y لهما نفس المتوسط الحسابي ونفس الوسيط ولكنهما يختلفان في درجة التشتت حيث يتضح أن المتغير x أقل تشتتاً من المتغير y .

مما سبق يتبين أن التشتت هو مقياس لقوة تجمع البيانات حول بعضها أو اختلافها عن بعضها ، فإذا كانت القيم متباعدة فإن مقياس التشتت يكون كبيراً أما إذا كانت القيم متقاربة من بعضها فإن مقياس التشتت يكون صغيراً ، وهناك عدة مقاييس تصلح لمقياس درجة التشتت أهمها المدى ، الانحراف الربيعي ، الانحراف المتوسط ، والانحراف المعياري وسنقوم بدراسة كل من هذه المقاييس مع توضيح كيفية حسابها واستخداماتها .

1.3 المدى (Range)

يعرف المدى على أنه الفرق بين أقل قيمة وأكبر قيمة في المجموعة . ويعتبر من أبسط مقاييس التشتت وأسهلها حساباً ، غير أنه يعاب عليه اعتماده على القيمتين الطرفيتين فقط واللتين كثيراً ما تكونا شاذتين عن قيم المجموعة ، فإذا كانت القيمتين الطرفيتين من القيم الشاذة وكانت أحدهما كبيرة جداً والأخرى صغيرة جداً فإن المدى سوف يبالغ في إظهار تشتت هذه القيم ويظهره أكثر من حقيقته .

المدى = أكبر قيمة - أصغر قيمة (Range = Max. Value - Min. Value)

أما في حالة البيانات المبوبة فإن المدى يحسب كما يلي :

المدى = الحد الأعلى للفتة الأخيرة - الحد الأدنى للفتة الأولى

ويكون المدى مضللاً أيضاً في حالة مقارنة المجموعات التي يختلف عدد مفرداتها اختلافاً كبيراً هذا بالإضافة إلى صعوبة حسابه من الجداول التكرارية وخاصة إذا كانت مفتوحة .

مثال (1-3)

أوجد المدى للقيم التالية :

93 ، 70 ، 35 ، 65 ، 40 ، 99 ، 55

الحل :

المدى = أعلى قيمة - أصغر قيمة

99 - 35

= 64 درجة .

مثال (2-3)

الجدول (1-3) يبين ثلاثة مجموعات من البيانات . والمطلوب إجراء مقارنة بين تشتت هذه المجموعات باستخدام المدى .

جدول (1-3)

6	6	6	6	6	المجموعة الأولى
8	7	6	5	4	المجموعة الثانية
13	10	6	1	0	المجموعة الثالثة

الحل :

نلاحظ أن المجموعات الثلاثة لها نفس المتوسط الحسابي والوسيط ، فكليهما في جميع المجموعات يساوي القيمة 6 وذلك بالرغم من وجود اختلاف واضح بين المجموعات الثلاثة في انتشار القيم وتباعدها . ففي المجموعة الأولى نلاحظ عدم وجود اختلاف أو تشتت بين القيم فكل القيم متساوية ، وتساوي قيمة المتوسط الحسابي وهي القيمة 6 . أما في المجموعة الثانية فنلاحظ أن القيم تختلف عن بعضها البعض وعن متوسطها الحسابي ولكن ليس اختلافاً كبيراً ، أي أن تشتت القيم داخل هذه المجموعة صغير بينما في المجموعة الثالثة نلاحظ أن القيم تختلف عن بعضها البعض وعن متوسطها الحسابي اختلافاً كبيراً ، أي أن تشتت القيم داخل هذه المجموعة أكبر . بالنسبة لمدى المجموعات الثلاثة يكون كما يلي :

مدى المجموعة الأولى = $6 - 6 = 0$.

مدى المجموعة الثانية = $8 - 4 = 4$.

مدى المجموعة الثالثة = $13 - 0 = 13$.

بما أن مدى المجموعة الأولى هو صفر ، فذلك يعني أنه لا يوجد اختلاف أو تشتت بين قيم هذه المجموعة أي أن كل قيم هذه المجموعة متساوية ، أما المجموعتين الثانية والثالثة فنلاحظ أن مدى المجموعة الثالثة أكبر من مدى المجموعة الثانية وهذا يعني أن تشتت المجموعة الثالثة أكبر من تشتت الثانية .

مثال (3-3)

قارن بين تشتت درجات مادة الإحصاء لثلاث مجموعات من الطلبة المبينة في الجدول (2-3) .

جدول (2-3)

—	—	80	45	70	30	62	50	المجموعة الأولى
35	77	82	65	55	70	0	20	المجموعة الثانية
75	65	63	60	50	85	49	54	المجموعة الثالثة

الحل :

مدى المجموعة الأولى = $30 - 80 = 50$ درجة .

مدى المجموعة الثانية = $0 - 82 = 82$ درجة .

مدى المجموعة الثالثة = $49 - 85 = 36$ درجة .

نلاحظ أن المجموعة الثالثة أقل تشتتاً تليها المجموعة الأولى ثم الثالثة .

3.3 الانحراف الربيعي (The Quartile Deviation)

أشرنا إلى أن من عيوب المدى اعتماده على القيم الطرفية التي غالباً ما تكون متطرفة وشاذة ، ويمكن التغلب على هذا السبب بحذف بعض القيم الطرفية فإذا أهملنا الربع الأول والربع الأخير من هذه القيم فإنه يمكن

الحصول على مقياس للتشتت يعتبر أفضل من المدى ويعتمد في حسابه على كل من الربيعين الأدنى والأعلى ويسمى بالانحراف الربيعي وهو عبارة عن نصف المدى الربيعي أي أن :

$$\text{الانحراف الربيعي} = \frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{2}$$

أو بمعنى آخر :

$$Q.D. = \frac{R_3 - R_1}{2} \dots\dots\dots(1-3)$$

حيث أن :

Q.D - الانحراف الربيعي .

R₃ - الربيع الأعلى .

R₁ - الربيع الأدنى .

ويمكن الحصول عليه بيانياً وذلك برسم منحنى التكرار المتجمع الصاعد أو النازل واستخراج قيمة الربيعين من الرسم ثم حساب نصف المدى بينهما .

ولغرض استنتاج الانحراف الربيعي لدرجات اختبار الذكاء للخمسين مهندس العينة العشوائية التي تم دراستها في الباب الأول الجدول (1-6) أو الشكل (1-7) لغرض حساب قيمة الربيعين الربيع الأدنى والربيع الأعلى ، ومن الفقرة (2 - 9) في الباب السابق فقد تم مسبقاً حساب قيمة كل من الربيع الأدنى والربيع الأعلى لدرجات اختبار الذكاء للخمسين مهندس ووجد أنهما يساويان الآتي :

$$R_1 = 106.8$$

$$R_3 = 124.1$$

وعلى هذا الأساس فإن قيمة الانحراف الربيعي تساوي :

$$Q.D. = \frac{R_3 - R_1}{2} = \frac{124.1 - 106.8}{2} = \frac{17.3}{2} = 8.65$$

مثال (4-3)

أحسب الانحراف الربيعي للبيانات التالية التي تمثل درجات سبعة طلبة في امتحان مادة الفيزياء الجامعية :

34 , 20 , 39 , 25 , 41 , 30 , 22

الحل :

نرتب القيم تصاعدياً ومنها نستنتج قيمة الربيعين حيث :

$$41 \quad \boxed{39} \quad 34 \quad 30 \quad 25 \quad \boxed{22} \quad 20$$

$$R_3 \quad \quad \quad R_1$$

أن ترتيب الربيع الأدنى (R_1) يساوي :

$$\frac{n+1}{4} = \frac{1+7}{4} = \frac{8}{4} = 2$$

وذلك لأن عدد القيم يساوي 7 ، إذن قيمة الربيع الأدنى هي القيمة الثانية في الترتيب التصاعدي أي أن الربيع الأدنى يساوي 22 .

أما ترتيب الربيع الأعلى (R_3) فتساوي :

$$\frac{3(n+1)}{4} = \frac{3(1+7)}{4} = \frac{24}{4} = 6$$

إن قيمة الربيع الأعلى هي القيمة السادسة في الترتيب التصاعدي أي أن الربيع الأعلى يساوي 39 ومنه :

$$\therefore Q.D. = \frac{R_3 - R_1}{2} = \frac{39 - 22}{2} = \frac{17}{2} = 8.5$$

مثال (5-3)

أحسب الانحراف الربيعي للبيانات والمبينة في الجدول (3-3) .

جدول (3-3)

البيانات	التكرار	التكرار المتجمع الصاعد
50 -	1	1
56 -	6	7
62 -	12	19
68 -	15	34
74 -	22	56
80 -	10	66
86 -	6	72
92 - 98	4	76

الحل :

عدد القيم يساوي 76 ورتبة الربع الأدنى تساوي :

$$\frac{76}{4} = 19$$

$$\therefore R_1 = L + \frac{\frac{n}{4} - k_1}{k_2 - k_1} = 62 + \frac{19 - 7}{19 - 7} \times 6 = 62 + 6 = 68$$

أما رتبة الربع الأعلى فتساوي :

$$57 = \frac{(76)(3)}{4}$$

$$\begin{aligned}\therefore R_3 &= L + \frac{\frac{3n}{4} - k_1}{k_2 - k_1} \\ &= 80 + \frac{57 - 56}{66 - 56} \times 6 \\ &= 80 + \frac{1}{10} \times 6 = 80 + 0.6 = 80.6\end{aligned}$$

إن الانحراف الربيعي يساوي :

$$\frac{R_1 - R_3}{2} = \frac{68 - 80.6}{2} = \frac{12.6}{2} = 6.3$$

أن الانحراف الربيعي وإن كان أفضل من المدى كمقياس للتشتت إلا أنه يعتمد على قيمتين فقط وهما الربع الأدنى R_1 والربع الأعلى R_3 ولا يضع في اعتباره كل القيم المتضمنة ومن ثم كانت هناك حاجة إلى مقاييس أفضل سنقوم بدراستها في البنود اللاحقة من هذا الباب .

4.3 الانحراف المتوسط (The Mean Deviation)

يمكن دراسة درجة تشتت أي مجموعة من القيم تبعاً لقرب مفرداتها من أو بعدها عن المتوسط العام للمجموعة فإذا استخدمنا الوسط الحسابي نظراً لأنه أكثر المتوسطات حساسية للتغيرات في أي قيمة من قيم المجموعة نجد أن مجوع انحرافات القيم عن متوسطها الحسابي يساوي صفراً ، وذلك من خواص المتوسط الحسابي كما أشرنا إلى ذلك سابقاً ، مهما كانت درجة التشتت والسبب في ذلك هو وجود انحرافات موجبة وأخرى سالبة تلاشي بعضها البعض ، فلو استبعدنا الإشارات أي أخذنا القيمة العددية لهذه الانحرافات نجدها تكبر كلما زاد انتشار القيم وبعدت عن متوسط المجموعة وتقل كلما قل تشتت القيم . لذلك يمكننا أن نأخذ متوسط القيمة العددية لانحرافات القيم عن متوسطها الحسابي كمقياس للتشتت وهذا المقياس يعرف بالانحراف المتوسط ويرمز له (M.D.) .

1.4.3 الخواص العامة للمتوسط (Properties of Mean Deviation)

- (a) يأخذ في حسابه كل القيم .
- (b) يمكن حسابه عن طريق الانحرافات .
- (c) لا يمكن حسابه من الجداول المفتوحة .
- (d) يتأثر بالقيم المتطرفة .

2.4.3 طرق حساب الانحراف المتوسط

أولاً : في حالة البيانات المفردة

يحسب الانحراف المتوسط المطلق في هذه الحالة بعد حساب المتوسط الحسابي ، ومن ثم نطرح كل القيم من المتوسط الحسابي ونأخذ القيمة المطلقة أي الموجبة للانحراف وبعد جمع الانحرافات المطلقة نقسمها على عدد القيم فنحصل على الانحراف المتوسط أي أن :

$$M.D. = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} \dots\dots\dots (2-3)$$

حيث أن :

$|x_i - \bar{x}|$ - هي القيمة العددية المطلقة لانحرافات القيم عن وسطها الحسابي .
 n - هي عدد المفردات .

فإذا توفرت لدينا القيم التالية مثلاً :

26 , 25 , 20 , 17 , 12

فإن المتوسط الحسابي لهذه القيم هو :

$$\bar{x} = \frac{12+17+20+25+26}{5} = \frac{100}{5} = 20$$

إن انحرافات القيم عن متوسطها الحسابي هو :

6 , 5 , 0 , 3 - , 8 -

$$\therefore \sum [x - \bar{x}] = 8 + 3 + 0 + 5 + 6 = 22$$

أي أن الانحراف المتوسط يساوي :

$$\frac{22}{5} = 4.4$$

مثال (3-6)

أوجد الانحراف المتوسط للأعداد :

77 , 85 , 63 , 45 , 70

الحل:

نجد أولاً المتوسط الحسابي للقيم ويساوي :

$$\bar{x} = \frac{45 + 63 + 70 + 77 + 85}{5} = \frac{340}{5} = 68$$

ثم بترتيب القيم والحدود كما يبين الجدول (4-3) حيث

جدول (4-3)

القيمة المطلقة	الانحراف عن الوسط الحسابي	الوسط الحسابي \bar{x}	القيمة
$23 = 23 - $	$23 - = 68 - 45$	68	45
$5 = 5 - $	$5 - = 68 - 63$	68	63
$2 = 2 - $	$2 - 68 - 70$	68	70
$9 = 9 - $	$9 = 68 - 77$	68	77
$17 = 17 - $	$17 = 68 - 85$	68	85
56	—————	Total	المجموع

$$\therefore M . D . = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} = \frac{56}{5} = 11.2$$

ثانياً : في حالة البيانات المبوبة

لإيجاد الانحراف المتوسط في حالة التوزيعات التكرارية تتبع الخطوات التالية :

(a) يتم تحديد مراكز الفئات أو الفترات \bar{x}_i .

(b) يحسب الوسط الحسابي كما في السابق من المعادلة $\bar{x} = \frac{\sum \bar{x}_i f_i}{\sum f_i}$

(c) تستخدم المعادلة التالية لحساب الانحراف المتوسط :

$$M.D. = \frac{\sum_{i=1}^n f_i |\bar{x}_i - \bar{x}|}{\sum_{i=1}^n f_i} \quad \dots\dots\dots (3-3)$$

إن الجدول (3-5) يبين حساب الانحراف المتوسط لدرجات اختبار الذكاء للخمسين مهندس الذي تم حساب المتوسط الحسابي له مسبقاً في الجدول رقم (1-2) في الباب الثاني وكانت قيمته 116 .

جدول (3 - 5)

إيجاد الانحراف المتوسط لدرجات الذكاء للخمسين مهندس

الفئات x_i	تكرار الفئة f_i	مركز الفئة \bar{x}_i	$ \bar{x}_i - \bar{x} $	$f_i \bar{x}_i - \bar{x} $
- 90	3	95	21	63
- 100	14	105	11	154
- 110	16	115	1	16
- 120	11	125	9	99
- 130	4	135	19	76
150- 140	2	145	29	78
المجموع	50	—	—	466

ولحساب الانحراف المتوسط من الجدول التكراري (5-3) نوجد الوسط الحسابي كما ذكرنا سابقاً ، بإحدى الطرق التي تمت مناقشتها في الباب السابق ثم نوجد القيمة العددية للانحرافات عن المتوسط الحسابي $|\bar{x}_i - \bar{x}|$ وأخيراً نضرب كل قيمة من هذه القيم في التكرار المناظر لنحصل على $f_i |\bar{x}_i - \bar{x}|$ ثم نجمع حاصل الضرب وبالقسمة على مجموع التكرارات نحصل على الانحراف المتوسط بالشكل التالي :

$$M .D. = \frac{\sum_{i=1}^n f_i |\bar{x}_i - \bar{x}|}{\sum_{i=1}^n f_i} = \frac{466}{50} = 9.32$$

مثال (7-3)

أحسب الانحراف المتوسط لدرجات 50 طالب والمبينة في الجدول (6-3) :

جدول (6-3)

الفئات	-0	-10	- 20	- 30	- 40	50 - 60
التكرار	3	10	12	20	4	1

الحل :

نرتب الجدول (6-3) بالشكل الموضح في الجدول (7-3) ونوجد كل من المتوسط الحسابي أولاً ثم الانحراف المتوسط ثانياً حيث :

جدول (7-3)

الفئة x	التكرار f_i	المركز \bar{x}_i	$\bar{x}_i f_i$	$\bar{x}_i - \bar{x}$	$ \bar{x}_i - \bar{x} $	$f_i \bar{x}_i - \bar{x} $
0 -	3	5	15	23 -	23	69
10 -	10	15	150	13 -	13	130
20 -	12	25	300	3 -	3	36
30 -	20	35	700	7	7	140
40 -	4	45	180	17	17	68
50 - 60	1	55	55	27	27	27
المجموع	50	—	1400	—	—	470

من الجدول نجد أن المتوسط الحسابي يساوي :

$$\bar{X} = \frac{\sum \bar{x}_i f_i}{\sum f_i} = \frac{1400}{50} = 28$$

إن الانحراف المتوسط يساوي :

$$M.D = \frac{\sum f_i |\bar{x}_i - \bar{x}|}{\sum f_i} = \frac{470}{50} = 9.4$$

5.3 الانحراف المعياري (The Standard Deviation δ)

وهو من أهم مقاييس التشتت وأكثرها استخداماً وذلك لدخوله في حساب الكثير من المقاييس الإحصائية الأخرى . وهو مثل الانحراف المتوسط في اعتماده على كل قيم المجموعة ، ونحصل عليه بتربيع انحرافات القيم عن وسطها الحسابي بدلاً من إهمال الإشارات كما في حالة الانحراف المتوسط وبذلك نحصل على :

$$s^2 = \frac{\sum (x - \bar{x})^2}{n} \dots\dots\dots (4-3)$$

وهذه الصيغة تعطي لنا ما يسمى بالتباين (Variance) ، وهو عبارة عن متوسط مربعات انحرافات القيم عن متوسطها الحسابي ولكي نحصل على مقياس للتشتت يكون بنفس وحدات المتغير x نأخذ الجذر التربيعي فنحصل على الانحراف المعياري من المعادلة التالية :

$$s = \sqrt{\frac{1}{n} \sum (x - \bar{x})^2} \dots\dots\dots (5-3)$$

أي أن الانحراف المعياري (S) هو الجذر التربيعي للتباين (S^2) ، ومن الواضح أن الحساب بالطريقة الحسابية يحتاج لحسابات كثيرة وخاصة إذا كثر عدد تلك المفردات (n) وإذا أحتوى المتوسط الحسابي (\bar{x}) على كسور مما ينتج عنه احتواء الانحرافات على كسور أيضاً ومن ثم صعوبة حساب مربعاتها لذلك كان من الأفضل استخدام صيغة أخرى لحساب التباين لا تتضمن هذه العمليات الحسابية كما يلي :

$$\begin{aligned} s^2 &= \frac{\sum (x - \bar{x})^2}{n} = \frac{1}{n} \sum (x - \bar{x})^2 \\ &= \frac{1}{n} \sum (x^2 - 2x\bar{x} + \bar{x}^2) \\ &= \frac{1}{n} (\sum x^2 - \sum 2x\bar{x} + \sum \bar{x}^2) \\ &= \frac{\sum x^2}{n} - 2\bar{x} \frac{\sum x}{n} + \frac{n\bar{x}^2}{n} = \frac{\sum x^2}{n} - 2\bar{x} \frac{\sum x}{n} + \bar{x}^2 \end{aligned}$$

وبما أن :

$$\bar{x} = \frac{\sum x}{n}$$

أنن :

$$\begin{aligned} S^2 &= \frac{\sum x^2}{n} - 2 \left(\frac{\sum x}{n} \right) \left(\frac{\sum x}{n} \right) + \left(\frac{\sum x}{n} \right)^2 \\ &= \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2 \dots\dots\dots (6-3) \end{aligned}$$

أي أن التباين هو متوسط المربعات ناقص مربع المتوسط ، ونستنتج من ذلك صيغة سهلة لحساب الانحراف المعياري وهي :

$$S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2} \dots\dots\dots (7-3)$$

وإذا أردنا تسهيل العمليات الحسابية أكثر من ذلك يمكن اختيار وسط فرضي من بين القيم ونحسب الانحرافات عن هذا الوسط الفرضي . فإذا كان هذا الوسط الفرضي هو (a) نجد أن التباين :

$$S^2 = \frac{1}{n} \sum [(x-a) - (\bar{x}-a)]^2 \dots\dots\dots (8-3)$$

وهنا نجد أن :

$x - a = d$ هي انحرافات القيم عن الوسط الفرضي a ومن دراستنا للمتوسط الحسابي سبق أن رأينا أن :

$$\bar{x} = a + \frac{\sum d}{n} = a + \bar{d} \quad \dots\dots\dots(9-3)$$

$$\therefore (\bar{x} - a) = \bar{d} \quad \dots\dots\dots(10-3)$$

$$\therefore S^2 = \frac{1}{n} \sum (d - \bar{d})^2 \dots\dots\dots(11-3)$$

أي أنه يمكن أن نستبدل قيم المفردات (x) بانحرافات هذه القيم عن الوسط الفرضي دون أن يؤثر ذلك على حساب التباين والانحراف المعياري فنجد أن :

$$\delta = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum \bar{d}}{n}\right)^2} \quad \dots\dots\dots(12-3)$$

ويسمى استخدام الطريقة الأخيرة بالصيغة المختصرة (Shortcut Formula) بينما تعرف الصيغة السابقة بالطريقة المطولة (Prolongated Method) وفيما يلي استخدام الطريقتين لحساب الانحراف المعياري للبيانات التالية :

26 , 25 , 20 , 17 , 12

أولاً بالطريقة المطولة :

$$\sum x = 12 + 17 + 20 + 25 + 26 = 100$$

$$\sum x^2 = 144 + 289 + 400 + 625 + 676 = 2134$$

$$\therefore S^2 = -\sqrt{\left(\frac{2134}{5}\right) - \left(\frac{100}{5}\right)^2} = \sqrt{426.8 - 400} = \sqrt{26.8} = 5.177$$

ثانياً بالطريقة المختصرة :

بما أن $n = 5$ وبأخذ القيمة 17 كوسط فرضي نجد أن انحرافات القيم عن الوسط الفرضي هي - 5 , 0 , 3 , 8 , 9 لذلك :

$$\sum d = -5 + 0 + 3 + 8 + 9 = 15$$

$$\sum d^2 = 25 + 0 + 9 + 64 + 81 = 179$$

$$\therefore S = \sqrt{\left(\frac{179}{5}\right) - \left(\frac{15}{5}\right)^2} = \sqrt{35.8 - 9} = \sqrt{26.8} = 5.177$$

ويبين المثال التالي بعض الخصائص العامة للانحراف المعياري التي تساعد في تسهيل العمليات الحسابية :

يحتوي العمود الأول على عشرة من القيم وقد أشتقت من هذه القيم الأصلية قيم أخرى مرة بإضافة رقم ثابت (10 مثلاً) ومرة أخرى بطرح الرقم الثابت ومرة أخرى بالضرب في الرقم الثابت وفي العمود الأخير بالقسمة على الرقم الثابت كما هو مبين في الجدول (8-3) :

جدول (8-3)

$x + 10$	$(x) (10)$	$x - 10$	$x + 10$	x
0.1	10	6 -	11	1
0.5	50	5 -	15	5
0.7	70	3 -	17	7
0.8	80	2 -	18	8
1.2	120	2	22	12
1.5	150	5	25	15
1.6	160	6	26	16
2.0	200	10	30	20
2.2	220	12	32	22
3.4	340	24	44	34
14.0	1400	40	240	140
28.04	280400	1004	6604	2804
1.4	140	4	24	14
0.844	8440	84.4	84.4	84.4
0.9187	9187	9.187	9.187	9.187

مجموع القيم
مجموع مربعات القيم
مكوسط القيم
الكتباين
الانحراف المعياري

يتضح من النتائج السابقة المبينة أسفل الجدول (3-8) أن كلاً من التباين والانحراف المعياري لا يتأثران بإضافة أو طرح رقم ثابت من القيم الأصلية ، بينما يتأثران بعمليات الضرب والقسمة فقط فإذا قسمت القيم الأصلية على رقم ثابت ، فإنه لا بد وأن نضرب تباين القيم المختصرة في مربع الرقم الثابت وأن نضرب الانحراف المعياري للقيم المختصرة في الرقم الثابت وذلك لنستنتج التباين والانحراف المعياري للقيم الأصلية .

فمثلاً لكي نستنتج التباين للقيم الأصلية يجب ضرب قيمة التباين المحسوب في العمود الأخير في 100 . ولكي نستنتج الانحراف المعياري للقيم الأصلية يجب ضرب قيمة الانحراف المعياري للقيم المختصرة في العمود الأخير في 10 .

6.3 إيجاد الانحراف المعياري للبيانات المبوبة

(Standard Deviation for Data Set)

لإيجاد الانحراف المعياري للبيانات المبوبة في جداول تكرارية تصبح صيغة التباين بالشكل التالي :

$$S^2 = \frac{1}{n} \sum f(x - \bar{x})^2 \dots\dots\dots(13-3)$$

حيث :

$$\sum f = n = \text{عدد المفردات}$$

\bar{x} : هي مراكز الفئات ولصعوبة استخدام هذه الصيغة يمكن وضعها على الصورة التالية :

$$S = \sqrt{\frac{\sum \bar{x}_i^2 f_i}{\sum f_i} - \left(\frac{\sum \bar{x}_i f_i}{\sum f_i} \right)^2} \dots\dots\dots(14-3)$$

وتسمى طريقة الانحراف المعياري باستخدام هذه الصيغة بالطريقة المطولة .

جدول (3 - 9)

إيجاد الانحراف المعياري لدرجات الذكاء للخمسين مهندس

الفئات x_i	التكرار f_i	مركز الفئة \bar{x}	$\bar{x}_i f_i$	$\bar{x}_i^2 f_i$
- 90	3	95	285	27075
- 100	14	105	1470	154350
- 110	16	115	1840	211600
- 120	11	125	1375	171875
- 130	4	135	540	72900
150- 140	2	145	290	42050
المجموع	50	—	5800	679850

ولحساب الانحراف المعياري لدرجات الذكاء للخمسين مهندس باستخدام البيانات المدرجة في الجدول (3-3) ، نجد كل من التباين والانحراف المعياري كما يلي :

$$S^2 = \frac{679850}{50} - \left(\frac{5800}{50} \right)^2 = 13597 - 13456 = 141$$

$$\therefore S = \sqrt{141} = 11.874$$

إن التباين يساوي 141 والانحراف المعياري يساوي 11.874 .

7.3 إيجاد الانحراف المعياري بطريقة الانحرافات (Standard Deviation Using Divergence Method)

لتسهيل العمليات الحسابية يمكن اختيار وسط فرضي من بين مراكز الفئات وتحسب الانحرافات عن هذا الوسط الفرضي وفي هذه الحالة يمكن كتابة الصيغة السابقة على الشكل التالي :

$$S = \sqrt{\frac{\sum d_i^2 f_i}{\sum f_i} - \left(\frac{\sum d_i f_i}{\sum f_i} \right)^2} \dots\dots\dots(15-3)$$

واضح أن هذه العلاقة أفضل وأسهل كثيراً من العلاقة (14-3) ، لذلك يجب استخدامها دائماً لتسهيل العمليات الحسابية . ويلاحظ أنه عند استخدام هذه الصيغة تتبع نفس الخطوات التي سبق شرحها عند حساب الوسط الحسابي ثم نزيد على الجدول عموداً واحداً هو ($d_i^2 f_i$) كما يتضح من الجدول (10-3) ، وتسمى هذه الطريقة " طريقة الانحرافات " نظراً لاعتمادها على حساب الانحرافات عن الوسط الفرضي الذي نختاره لتسهيل العمليات الحسابية .

جدول (10-3)

إيجاد الانحراف المعياري لدرجات الذكاء للخمسين مهندس باستخدام وسط فرضي

الفئات	التكرار f_i	مركز الفئة \bar{x}_i	الانحراف d_i	$d_i f_i$	$d_i^2 f_i$
90	3	95	- 20	- 60	1200
- 100	14	105	- 10	- 140	1400
- 110	16	115	صفر	صفر	صفر
- 120	11	125	10	110	1100
- 130	4	135	20	80	1600
150 - 140	2	145	30	60	1800
المجموع	50	—	—	50	7100

ولحساب الانحراف المعياري لدرجات الذكاء بطريقة الانحرافات اخترنا القيمة 115 كوسط فرضي من بين مراكز الفئات ، ثم حسبنا الانحرافات عنه كما يبين الجدول (10-3) حيث نجد أن :

$$S = \sqrt{\frac{7100}{50} - \left(\frac{50}{50}\right)^2} = \sqrt{142 - 1} = \sqrt{141} = 11.874$$

8.3 إيجاد الانحراف المعياري بطريقة الانحرافات المختصرة (Standard Deviation by Shortcut Deviations Method)

في حالة التوزيعات التكرارية المنتظمة أي الجداول التكرارية ذات الفئات المتساوية الطول يمكننا استخدام الانحرافات المختصرة وذلك بقسمة انحرافات القيم عن الوسط الفرضي على طول الفئة أي أن :

$$\frac{d_i}{L} = \bar{d}_i$$

حيث أن :

\bar{d}_i - هو الانحراف المختصر .

L - هو طول الفئة .

بذلك نجد أن العلاقة (15-3) لحساب الانحراف المعياري تصبح على الصورة التالية :

$$S = L \sqrt{\frac{\sum \bar{d}_i^2 f_i}{\sum f_i} - \left(\frac{\sum \bar{d}_i f_i}{\sum f_i}\right)^2} \dots\dots\dots(16-3)$$

وهذه العلاقة وإن كانت أسهل من العلاقة السابقة إلا أنها لا يمكن أن تستخدم إلا في حالات التوزيعات المنتظمة ، أي في حالة تساوي أطوال الفئات .

جدول (11-3)

إيجاد الانحراف المعياري لدرجات الذكاء لخمسين مهنس باستخدام طريقة
الانحرافات المختصرة

التفاه	ال تكرار f_i	مركز الفئة \bar{x}_i	الانحراف d_i	\bar{d}_i	$\bar{d}_i f_i$	$\bar{d}_i^2 f_i$
- 90	3	95	- 20	- 2	- 6	12
- 100	14	105	- 10	- 1	- 14	14
- 110	16	115	صفر	صفر	صفر	صفر
- 120	11	125	10	1	11	11
- 130	4	135	20	2	8	16
150-140	2	145	30	3	6	18
لمجموع	50	—	—	—	5	71

ولحساب الانحراف المعياري للمثال السابق باستخدام الانحرافات
المختصرة نتبع نفس الخطوات التي سبق دراستها عند حساب المتوسط
الحسابي ثم نضيف عموداً آخر هو $(\bar{d}_i f_i)$ كما موضح في الجدول (11-3)
ثم نوجد قيمة الانحراف المعياري حيث :

$$\begin{aligned}
 S &= 10 \sqrt{\frac{71}{50} - \left(\frac{5}{50}\right)^2} = \sqrt{1.42 - 0.01} \\
 &= \sqrt{1.41} \\
 &= 11.874
 \end{aligned}$$

مثال (8-3)

أحسب الانحراف المعياري لدرجات 70 طالب في مادة الإحصاء والمبينة في الجدول (12-3).

جدول (12-3)

الفئات	-10	-15	-20	-25	-30	-35	-40	-45	-50	-55	65-60
التكرار	1	3	7	13	15	10	8	6	4	2	1

الحل :

عدد الفئات هو فردي لذلك فإن فئة الوسط الفرضي هي (40-35) ، أي أن الوسط الفرضي هو مركز تلك الفئة = 37.5 ومن الجدول (12-3) نجد أن $n = 70$ وطول الفئة = 5 ، ثم نحسب الانحرافات كما يبين الجدول (13-3) :

جدول (13-3)

الفئات	التكرار f_i	مركز الفئة	الانحراف d_i	معدل الانحراف \bar{d}_i	$\bar{d}_i f_i$	$\bar{d}_i^2 f_i$
-10	1	12.5	25-	5-	5-	25
-15	3	17.5	20-	4-	12-	48
-20	7	22.5	15-	3-	21-	63
-25	13	27.5	10-	2-	26-	52
-30	15	32.5	5-	1-	15-	15
-35	10	37.5	0	0	0	0
-40	8	42.5	5	1	8	8
-45	6	47.5	10	2	12	24
-50	4	52.5	15	3	12	36
-55	2	57.5	20	4	8	32
65-60	1	62.5	25	5	5	25
المجموع	70	—	—	—	34 -	328

$$\therefore S = (5) \sqrt{\frac{328}{70} - \left(\frac{-34}{70}\right)^2} = (5) \sqrt{4.686 - (0.486)^2}$$

$$= (5) \sqrt{4.686 - 0.2362} = (5)(2.11) = 10.553$$

مثال (9-3) :

أوجد كل من الانحراف المتوسط والانحراف الربيعي والانحراف المعياري للتوزيع التكراري وكما هو مبين في الجدول (14-3) :

جدول (14-3)

الفئات	- 20	- 30	- 40	- 50	- 60	70 وأقل من 80
التكرار	10	15	30	22	14	9

الحل:

أولاً لإيجاد الانحراف المتوسط نحسب المتوسط الحسابي وننظم الجدول (14-3) على النحو المبين في الجدول (15-3) :

جدول (15-3)

الفئات	التكرار f_i	مركز الفئة \bar{x}_i	$\bar{x}_i f_i$	$ \bar{x}_i - \bar{x} $	$f_i \bar{x}_i - \bar{x} $
20-	10	25	250	24.2	242
30-	15	35	525	14.2	213
40-	30	45	1350	4.2	126
50-	22	55	1210	5.8	127.6
60-	14	65	910	15.8	221.2
80- 70	9	75	675	25.8	232.2
المجموع	100	—	4920	—	1162

بما أن المتوسط الحسابي للبيانات :

$$\bar{x} = \frac{4920}{100} = 49.2$$

إن الانحراف المتوسط

$$M.D. = \frac{1162}{100} = 11.62$$

ثانياً لإيجاد الانحراف الربيعي نحسب كلاً من الربع الأدنى والأعلى وهذا يتطلب تكوين جدول التكرار المتجمع الصاعد كما يلي:

نجد ترتيب الربع الأدنى حيث :

$$\frac{100}{4} = 25$$

وحيث أن هذه القيمة موجودة بجدول التكرار المتجمع الصاعد لذلك فإن الربع الأدنى يساوي 40 ، وهي القيمة المقابلة لترتيب الربع بالجدول (3-16) :

جدول (3-16)

حدود الفئات	التكرار المتجمع الصاعد
أقل من 20	صفر
أقل من 30	10
أقل من 40	25
أقل من 50	55
أقل من 60	77
أقل من 70	91
أقل من 80	100

ترتيب الربع الأعلى يساوي : $100 \times 3/4 = 75$

وينضح من جدول التكرار المتجمع الصاعد (3-16) ، أن فئة الربع الأعلى هي (50 وأقل من 60) والتكرار المتجمع الصاعد السابق لترتيب الربع الأعلى يساوي 55 والتكرار المتجمع الصاعد اللاحق يساوي 77 وعلى هذا الأساس فإن الربع الأعلى :

$$R_3 = 50 + \left(\frac{75-55}{77-55} \right) (10) = 50 + \frac{(20)(10)}{22} = 50 + 9.1 = 59.1$$

أي أن الانحراف الربيعي :

$$\frac{R_1 - R_3}{2} = \frac{19.1}{2} = \frac{40 - 59.1}{2} = 9.55$$

ثالثاً لإيجاد الانحراف المعياري باستخدام طريقة الانحرافات المختصرة نكون الجدول (3-17) ، ثم نختار للبيانات وسطاً فرضياً وبما أن عدد الفئات هو 6 أي زوجي لذلك قيمة الوسط الفرضي هي مركز الفئة المقابل لأعلى تكرار ويساوي 45 .

جدول (3-17)

الفئات	التكرار	مركز الفئة	الانحرافات d_i	الانحرافات المختصرة \bar{d}_i	$\bar{d}_i f_i$	$\bar{d}_i^2 f_i$
- 20	10	25	20 -	2 -	20 -	40
- 30	15	35	10 -	1 -	15 -	15
- 40	30	45	0	0	0	0
- 50	22	55	10	1	22	22
- 60	14	65	20	2	28	56
80 - 70	9	75	30	3	27	81
المجموع	100	—	—	—	42	214

$$\therefore S = (10) \sqrt{\frac{214}{100} - \left(\frac{42}{100}\right)^2} = (10) \sqrt{2.14 - (0.42)^2}$$

$$= (10) \sqrt{2.14 - 0.1764} = (10)(1.4013) = 14.013$$

9.3 معامل الاختلاف (Coefficient of Variation)

عند مقارنة التوزيعات التكرارية تقابلنا عادةً صعوبة وهي الاختلاف في وحدة القياس . فإذا أردنا مثلاً مقارنة تشتت أطوال مجموعة من الأشخاص بتشتت توزيع أوزانهم نجد أن وحدات القياس المستخدمة في الحالة الأولى هي السنتيمترات ، بينما الوحدات المستخدمة في الحالة الثانية هي الكيلوجرامات وللتخلص من هذه المشكلة يمكن استخدام مقياس نسبي للتشتت لا يتأثر بوحدات القياس المستخدمة في كل من التوزيعين . فإذا قسمنا الانحراف المعياري لكل توزيع على الوسط الحسابي لنفس التوزيع نحصل على مقياس نسبي للتشتت يعرف بمعامل الاختلاف D حيث أن :

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times 100$$

$$\therefore D = \frac{S}{\bar{x}} \times 100 \dots\dots\dots(17-3)$$

واستخدام التشتت النسبي لا يقتصر فقط على التخلص من وحدات القياس ولكن أيضاً لمقارنة التوزيعات التي يوجد فرق كبير بين متوسطاتها . حتى ولو كانت مقاسة بنفس وحدات القياس . فمثلاً إذا كان المتوسط الحسابي لقوة التحمل لنوع معين من الأسلاك هو 128.64 باوند بانحراف معياري 15.37 باوند ، والمتوسط الحسابي لقوة التحمل لنوع آخر من الأسلاك هو 87.66 باوند بانحراف معياري 14.12 باوند ، وأردنا مقارنة درجة التشتت

للتنوعين من الأسلاك فلا يمكن مقارنة القيم المطلقة لتشتت النوعين نظراً لاختلاف متوسطيهما . لذلك نصب معامل الاختلاف لكل من النوعين حيث :

معامل الاختلاف للنوع الأول يساوي :

$$\frac{15.37}{128.64} \times 100 = 11.95 \%$$

أما معامل الاختلاف للنوع الثاني فيساوي :

$$\frac{14.12}{87.66} \times 100 = 16.11 \%$$

وهكذا نجد أن التشتت النسبي مقاساً بمعامل الاختلاف يشير إلى أن النوع الأول أقل تشتتاً من النوع الثاني ، وهذه النتيجة التي نحصل عليها تخالف تماماً النتيجة التي نحصل عليها بالاعتماد على المقياس المطلق للتشتت الانحراف المعياري فقط .

وفي حالة الجداول التكرارية المفتوحة لا يمكن حساب كل من المتوسط الحسابي والانحراف المعياري . لذلك تستخدم صيغة أخرى تعتمد على الربيعين الأعلى والأدنى . ولما كان معامل الاختلاف عبارة عن مقياس للتشتت مقسوماً على مقياس المتوسط فإنه يمكن إيجاده بقسمة الانحراف الربيعي على الوسيط وباعتبار أن الوسيط يساوي الوسط الحسابي للربيعين :

$$\therefore \text{معامل الاختلاف} = \frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{\text{الربيع الأعلى} + \text{الربيع الأدنى}} \times 100 \quad \text{أو}$$

$$D = \frac{R_3 - R_1}{R_3 + R_1} \times 100 \dots\dots\dots (18-3)$$

وتستخدم هذه الصيغة أيضاً إذا أردنا إيجاد معامل الاختلاف بيانياً حيث يمكن حساب قيمة الربيعين الأعلى والأدنى من منحني التكرار المتجمع الصاعد.

10.3 الوحدات المعيارية (Standard Units z – Score)

إذا كان لدينا مجموعة من المفردات ثم حسبنا المتوسط الحسابي \bar{x} ، والانحراف المعياري S لهذه المجموعة ، ثم طرحنا قيمة المتوسط الحسابي من كل مفردة من مفردات المجموعة وقسمنا الناتج على قيمة الانحراف المعياري فإن القيم الجديدة التي نحصل عليها تكون مقياسة بوحدات تعرف بالوحدات المعيارية فإذا رمزنا للقيم الجديدة بالرمز z نجد أن :

$$z = \frac{x - \bar{x}}{S} \dots\dots\dots (19-3)$$

حيث أن الوسط الحسابي للقيم \bar{x} يساوي صفراً والانحراف المعياري لها يساوي الواحد الصحيح .

ونفيدنا الصيغة المعيارية (6-19) في أنها تمكننا من مقارنة قيم المجموعات المختلفة وذلك بتحويل الوحدات المستخدمة في كل مجموعة إلى وحدات معيارية وذلك باستخدام الوسط الحسابي والانحراف المعياري .

مثال (3-9)

قامت باحثة في مجال التغذية بدراسة كميات البروتين المتناولة من قبل عدد من الأشخاص البالغ عددهم مائة شخص . ووجدت أن المتوسط الحسابي لمأخذ البروتين لهؤلاء الأشخاص هو 77 gm ، وكان الانحراف المعياري لمأخذ البروتين للأشخاص هو 8.8 gm . أحسب الوحدة المعيارية لمأخذ بروتين يساوي 93 gm .

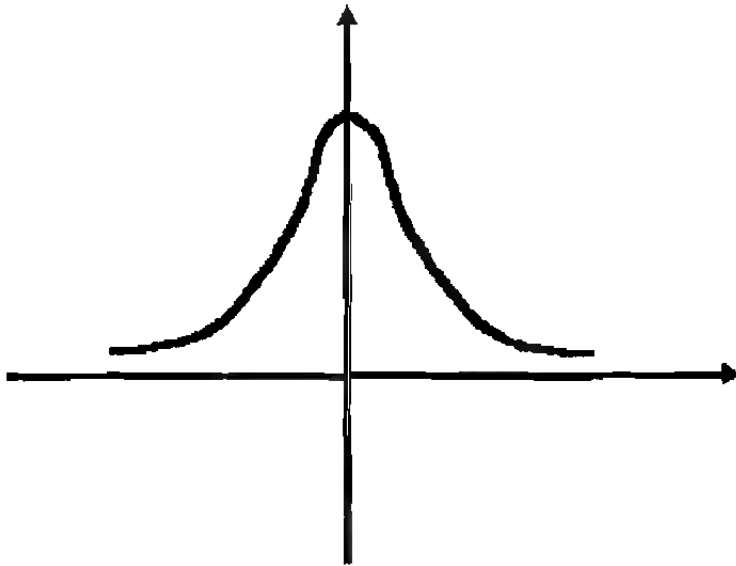
الحل :

نقوم بحساب الوحدة المعيارية بتطبيق العلاقة (3 - 19) نحصل على :

$$z = \frac{x - \bar{x}}{s} = \frac{93 - 77}{8} = \frac{16}{8} = 2$$

11.3 المنحنى التكراري المتماثل والتوزيعات الاعتدالية (Symmetric Distribution Curve & Moderate Distribution)

يكون التوزيع التكراري المتماثل متوازناً بحيث إن نصف المنحنى الذي يمثله يكون كل منهما صورة انعكاسية للآخر بالنسبة للخط الرأسى المار بقمة المنحنى كما موضح في الشكل (3- 1) .



الشكل (3- 1)

منحنى تكراري متماثل

وإذا كان المنحنى المتمثل يشبه الجرس (Bell – Shaped) وكانت معظم الحالات تتجمع حول مركز التوزيع وتقل تحت طرفي المنحنى وذلك بنسب معينة فإن التوزيع يسمى توزيعاً اعتدالياً . ومن الناحية النظرية فإن توزيع البيانات الخاصة بأية صفة من صفات مجتمع ما أو أي عينة غير متحيزة تتبع التوزيع الاعتدالي ، مثل توزيع درجات الذكاء للخمسين مهندس ومن الخواص الهامة في منحنى التوزيع التكراري أن :

المتوسط الحسابي (Arithmetic Mean) يساوي الوسيط (Median) ويساوي المنوال (Mode) ، كما أن المساحة الواقعة تحته تتوزع بنسب مرتبطة بالوسط الحسابي والانحراف المعياري للتوزيع .

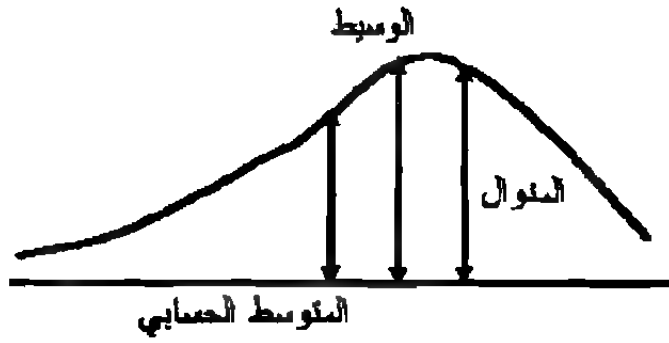
12.3 الحيود عن التوزيع الاعتدالي

(Diffraction from Moderate Distribution)

من الناحية العملية ، نجد أن الكثير من البيانات تحيد توزيعاتها عن الاعتدال في التوزيعات التكرارية ومنها :

(a) الالتواء (Skewness)

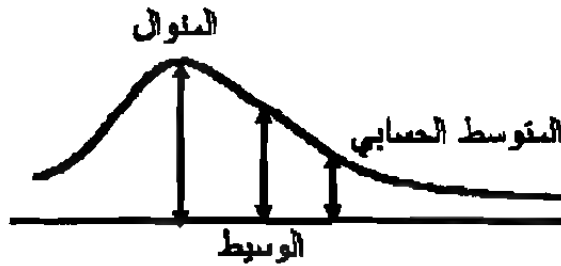
قد ينحرف التوزيع عن الاعتدال بحيث يصبح شكل المنحنى غير متماثل وذلك عندما تزداد القيم الكبرى مثلاً في التوزيع التكراري ، فيلتوي منحنى التوزيع تجاه اليسار حيث ينجذب المتوسط الحسابي تجاه القيم الشاذة الصغيرة ، بينما يكون المنوال والوسيط في صالح الأغلبية من الحالات ذات القيم الكبيرة كما موضح في الشكل (2-3) .



الشكل (3- 2)

المنحنى ذو الالتواء السالب

كذلك قد يزداد في التوزيع عدد القيم الصغرى فيلتوي منحنى التوزيع تجاه اليمين حيث ينجذب المتوسط الحسابي تجاه القيم الشاذة الكبيرة بينما يكون المتوال والوسيط في صالح الأغلبية من الحالات ذات القيم الصغيرة كما موضح في الشكل (3- 3) .



الشكل (3-3)

منحنى ذو التواء موجب

وهناك عدة مقاييس للالتواء نذكر منها المقياس التالي :

$$\text{معامل الالتواء} = \frac{3(\text{الوسط الحسابي} - \text{الوسيط})}{\text{الانحراف المعياري}}$$

$$Skewness Factor = \frac{3(\bar{x} - M)}{\delta} \dots\dots\dots(20-3)$$

ومن الواضح من هذه العلاقة أنه :

- 1- معامل الالتواء يساوي صفر في حالة المنحني المتماثل .
- 2- معامل الالتواء أكبر من صفر في حالة الالتواء تجاه القيم الكبرى .
- 3- معامل الالتواء أصغر من صفر في حالة الالتواء تجاه القيم الصغرى .

مثال (11-3)

أحسب معامل الالتواء للتوزيع التكراري المعطى في المثال (8-3) ثم أرسم شكل منحني التوزيع .

الحل:

ننظم جدول التكرار المتجمع الصاعد للفئات كما هو مبين في الجدول (18-3):

جدول (18-3)

الحدود العليا للفئات	التكرار المتجمع الصاعد
أقل من 10	0
أقل من 15	1
أقل من 20	4
أقل من 25	11
أقل من 30	24
أقل من 35	39
أقل من 40	49
أقل من 45	57
أقل من 50	63
أقل من 55	67
أقل من 60	69
أقل من 65	70

من الجدول (3-18) نلاحظ أن :

نجد رتبة الوسيط تساوي $35 = 70 / 2$ ، ومن ثم نجد معامل الالتواء حيث :

$$\begin{aligned}\therefore M &= 30 + \frac{35 - 24}{15} \times 5 \\ &= 30 + \frac{(5)(11)}{(15)} \\ &= 30 + \frac{11}{3} \\ &= 33.7\end{aligned}$$

أن الوسط الحسابي من المثال (3-8) يحسب كما يلي :

$$\bar{x} = 37.5 + \frac{(-34)(5)}{70} = 37.5 - 2.4 = 35.1$$

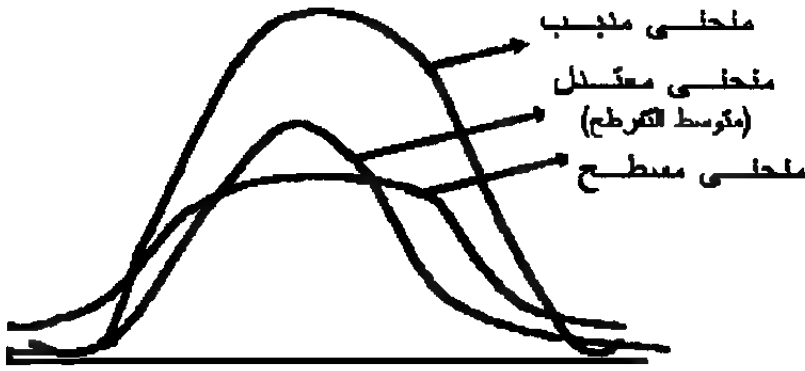
أما الانحراف المعياري من نفس المثال فيساوي $S = 10.5$ ، إذن معامل الالتواء يساوي :

$$0.4 = \frac{(33.7 - 35.1)(3)}{10.5}$$

وهذا يعني أن المنحني يعاني من التواء موجب بسيط . أما رسم شكل منحني التوزيع فنتركه كتمرين للطالب .

(b) التفريط (Oblation)

قد يكون شكل منحنى التوزيع متماثلاً ولكن التوزيع نفسه قد لا يكون اعتدالياً ، وذلك لاختلاف نسب التوزيع بين المركز والأطراف فقد يكون المركز ذا قيمة رفيعة حيث تزداد الحالات المتجمعة حول المركز وتزداد الحالات عند الطرفين . ويقاس ذلك بمقياس يسمى معامل التفريط . فإذا كان التوزيع ذا قمة رفيعة سمي التوزيع مدبب التفريط وإذا كان ذا قمة عريضة سمي التوزيع مسطح التفريط كما موضح في الشكل (4-3) .



الشكل (4-3)

وهناك عدة مقاييس للتفريط نذكر منها المقياس التالي :

$$\text{معامل التفريط} = \frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{2(\text{المئين التسعين} - \text{المئين العاشر})}$$

أو بالرموز :

$$\text{Oblation Factor} = \frac{R_3 - R_1}{2(P_{90} - P_{10})}$$

حيث أن :

- R_1 - الربع الأدنى .
- R_3 - الربع الأعلى .
- P_{10} - المئين العاشر .
- P_{90} - المئين التسعين .

وطبقاً لهذه المقاييس فإن منحنى التوزيع يكون :

- 1- اعتدالياً (Moderate) إذا كان معامل التفرطح يساوي 0.263 .
- 2- مسطحاً (Flatten) إذا كان معامل التفرطح أكبر من 0.263 .
- 2- مدبباً (Sharpened) إذا كان معامل التفرطح أقل من 0.263 .

مثال (3- 12)

أحسب معامل التفرطح في التوزيع المعطى في المثال (3- 8) .

الحل :

أن ترتيب P_{10} يساوي $70/10 = 7$ ومنه :

$$\begin{aligned}\therefore P_{10} &= 20 + \frac{7 - 4}{11 - 4} \times 5 \\ &= 20 + \frac{(5)(3)}{(7)} \\ &= 22.857\end{aligned}$$

أما ترتيب الربع الأدنى فهو :

$$R_1 = 70/4 = 17.5$$

$$\begin{aligned}\therefore R_1 &= 25 + \frac{17.5 - 11}{24 - 11} \times 5 \\ &= 25 + \frac{(5)(6.5)}{(13)} \\ &= 27.5\end{aligned}$$

أما ترتيب الربيع الأعلى فيساوي $R_3 = 3 \times 70 / 4$ ومنه :

$$\begin{aligned}\therefore R_3 &= 40 + \frac{52.5 - 49}{57 - 49} \times 5 \\ &= 40 + \frac{(5)(3.5)}{(8)} \\ &= 42.185\end{aligned}$$

أما ترتيب فيساوي :

$$\begin{aligned}P_{90} &= 70 \times 9 / 10 \\ &= 63\end{aligned}$$

إن $P_{90} - 50$ وعلى هذا الأساس فإن معامل التفرطح يساوي :

$$\begin{aligned}\therefore Oblation Factor &= \frac{R_3 - R_1}{2(P_{90} - P_{10})} \\ &= \frac{(42.185 - 27.5)}{2(50 - 22.857)} = 0.2702\end{aligned}$$

ومن هذا نرى أن لهذا التوزيع منحنى مسطحاً قليلاً بالنسبة لمنحنى التوزيع الاعتيادي .

13.3 تمرين

س1: أحسب كل من المدى والانحراف المتوسط والانحراف المعياري لكل مجموعة من البيانات ، وبعد ذلك أحسب معامل الاختلاف للمجموعة الثانية (b) :

a) 12 , 6 , 13 , 5 , 14 , 7 , 2

b) 11 , 1 , 2 , 9 , 6 , 4 , 7 , 3 , 5 , 2

س2: أوجد معامل الاختلاف للتوزيع التكراري المبين في الجدول (19-3) :

جدول (19-3)

الفاصل	أقل من 20	- 20	-25	- 30	35 وأقل من 40
التكرار	8	25	26	22	9

س3: الجدول (20-3) يبين التوزيع التكراري لمائة عامل من عمال أحد المصانع حسب فئات الأجر الأسبوعي بالدنانير .

جدول (20-3)

فئات الأجر	- 20	- 30	- 40	- 50	60 وأقل من 70
عدد العمال	26	40	18	10	3

المطلوب هو إيجاد كل من الربيع الأعلى والأدنى بيانياً ، ثم استنتاج معامل الاختلاف ، وحساب الانحراف المتوسط والانحراف المعياري .

س4: الجدول (21-3) يبين التوزيع التكراري لدخل 100 عائلة بالدنانير في الأسبوع .

جدول (21-3)

الفئات	- 5	- 10	- 15	- 20	- 25	- 30	- 35	40 - 45
التكرار	4	22	28	18	12	9	4	3

أوجد ما يلي :

- (a) حساب المتوسط الحسابي .
(b) حساب الانحراف المعياري .

س5: في عينة من 200 شخص في إحدى القرى وجد أن التوزيع التكراري للرجال المتزوجين حسب أعمارهم بالسنين كما يبينه الجدول (22-3) والمطلوب :

- (a) أوجد الانحراف المعياري والانحراف المتوسط للبيانات .
(b) أحسب معامل الاختلاف للبيانات .

جدول (22-3)

الفئات	- 15	- 25	- 35	- 45	- 55	65 وأقل من 75
التكرار	49	65	41	26	16	3

س6: حصل طالب على 84 درجة في أحد الامتحانات النهائية في مادة الاستاتيكا ، حيث كان متوسط درجات جميع الطلبة في هذه المادة 76 درجة بانحراف معياري 10 درجات ، وفي امتحان مادة خواص المواد حصل نفس الطالب على 90 درجة حيث كان متوسط درجات جميع الطلبة في هذه المادة 82 بانحراف معياري 16 درجة . في أي المادتين كان الطالب أكثر تفوقاً .

س7: أرسم منحنى التكرار المتجمع الصاعد للتوزيع المبين في الجدول (23-3) ثم أستنتج منه ما يلي :

(a) الانحراف المعياري .

(b) معامل الالتواء ومعامل التفرطح .

جدول (23-3)

الفئات	-50	-60	-70	-80	-90	-100	-110	-120	-130	150 - 140
التكرار	8	10	17	34	42	32	22	15	11	9

س8: أدرس التوزيع المبين في الجدول (24-3) ، إذا كان المنحنى متمائلاً أم غير متمائل ، ثم أحسب المتوسط الحسابي والانحراف المعياري والدرجة المعيارية لمراكز الفئات التي يشملها هذا التوزيع .

جدول (24-3)

الفئة	-50	-60	-70	-80	-90	-100	120 - 110
التكرار	10	15	30	50	25	20	10

س9: أوجد كلاً من التشتت المطلق والتشتت النسبي للتوزيع التكراري المبين في الجدول (25-3) :

جدول (25-3)

الفئات	أقل من 5	- 5	- 10	- 20	40 فأكثر
التكرار	8	32	50	7	3

وذلك باستخدام كل من :

(a) الطريقة البيانية .

(b) الطريقة الحسابية .

س10: أسفرت الدراسة التي أجريت عن الأجور الأسبوعية بالدنانير لعدد من العمال في مصنعين يعملان في صناعة واحدة عن النتائج الدرجة في الجدول (26-3) :

جدول (26-3)

المقياس	المصنع الأول	المصنع الثاني
الوسط الحسابي	62	54
الانحراف المعياري	16	15

المطلوب مقارنة تشتت الأجور الأسبوعية للعمال في المصنعين .

س11: إن الجدول التوزيع التكراري (27-3) يمثل درجات التحصيل في مادتي اللغة العربية والرياضيات لمائة طالب .

جدول (27-3)

النقطة	-0	-10	-20	-30	-40	-50	-60	-70	-80	90 - 100
اللغة العربية	1	3	2	13	29	31	13	2	4	2
الرياضيات	2	4	10	15	19	20	15	10	4	1

والمطلوب ما يلي :

1- قارن بين تشتتي الدرجات في التوزيعين عن طريق رسم مدرجين تكرارين لهما.

2- أحسب الوسط الحسابي والانحراف المعياري لكل من التوزيعين .

س12: قارن بين تشتتي المجموعتين A و B في كل من الحالات التالية المبينة في الجدول (28-3) باستخدام معامل الاختلاف مع التعليق على كل حالة من الحالات .

جدول (28-3)

المقياس	\bar{x}	S
A	100	10
B	50	4

المقياس	\bar{x}	S
A	100	5
B	120	5

المقياس	\bar{x}	S
A	65	5
B	65	10

س13: الجدول (29-3) يمثل توزيع المرتبات في إحدى السدائر الحكومية مقدرة بالدنانير لعند 140 مستخدم فيها :

جدول (29-3)

المرتبات	-70	-90	-110	-130	-150	-170	-190	210 - 230
التكرار	8	12	25	40	20	17	13	5

والمطلوب ما يلي :

(a) أحسب التشتت والانحراف الربيعي والمتوسط للبيانات .

(b) أحسب الانحراف المعياري للبيانات أعلاه .

(b) أحسب معامل الالتواء ومعامل التفرطح للبيانات .

الباب الرابع

الارتباط والانحدار

(Correlation and Regression)

- 1.4 مقدمة .
- 2.4 معامل الارتباط .
- 3.4 حساب معامل الارتباط .
- 4.4 إيجاد معامل الارتباط باستخدام الوسط الغرضي .
- 5.4 حساب معامل الارتباط للبيانات المبوبة .
- 6.4 معامل ارتباط الرتب سبيرمان .
- 7.4 معامل ارتباط الرتب سبيرمان في حالة الرتب التكرارية .
- 8.4 معامل الاقتران .
- 9.4 معامل التوافق .
- 10.4 الانحدار .
- 11.4 طريقة المربعات الصغرى .
- 12.4 معادلة خط انحدار x على y .
- 13.4 العلاقة بين معامل الارتباط ومعاملات الانحدار .
- 14.4 تمارين .

1.4 مقدمة

في الأبواب السابقة من هذا الكتاب تم عرض بعض المقاييس الإحصائية التي تصف متغير واحد منها المتوسط الحسابي والوسيط والمنوال والوسط الهندسي كمقاييس للنزعة المركزية ، والانحراف الربيعي والانحراف المعياري كمقاييس لتشتت القيم . أما في هذا الباب فسوف نقوم بدراسة العلاقة بين متغيرين أو أكثر بهدف معرفة الارتباط بين هذه المتغيرات .

ولدراسة الارتباط بين متغيرين نحتاج لمقياس يقيس لنا درجة العلاقة بينهما واتجاه هذه العلاقة ، فإذا وجدنا أن الزيادة في المتغير الأول تصاحبها زيادة في المتغير الثاني وأن النقص في المتغير الأول يصاحبه نقص في المتغير الثاني نقول أنه يوجد ارتباط طردي موجب بين هذين المتغيرين . أما لو كانت الزيادة في المتغير الأول يصاحبها نقص في المتغير الثاني وأن النقص في المتغير الأول تصاحبه زيادة في المتغير الثاني فأنا نقول أنه يوجد ارتباط عكسي أو سالب بين هذين المتغيرين .

وقد تقابلنا حالات نجد فيها أن الارتباط يكون تاماً ، سواء كان طردياً أو عكسياً ، وفي هذه الحالات نستطيع معرفة أحد المتغيرين لو عرفنا المتغير الآخر . والأمثلة على ذلك عديدة منها العلاقة بين مساحة الدائرة ونصف قطرها وطول ضلع المربع ومساحته وغيرها . وقد تقابلنا أيضاً حالات يعدم فيها الارتباط مثل دراسة العلاقة بين طول الفرد ودخله كما هو الحال في الكثير من الحالات الشائعة والتي تواجهنا كثيراً في الدراسات المختلفة والتي تتباين من الحالات التي يكون فيها الارتباط تاماً ولا يكون منعماً ولكن بين هذا وذاك . مثل دراسة العلاقة بين الطول والوزن أو العلاقة بين التقدير الذي حصل عليه بعض الطلبة في مادتين وغيرها وسوف نقوم بدراسة هذه الأمثلة بالتفصيل .

ويجب ملاحظته أن وجود ارتباط بين متغيرين لا يبين ما إذا كان أحدهما تابع للآخر فإذا كان لدينا متغيرين هما x ، y ووجدنا أن هناك بينهما ارتباطاً قوياً فإن هذا لا يوضح ما إذا كانت x تؤثر في y أو أن y تؤثر في x أم أن هناك عامل مشترك يؤثر في كل منهما وهو الذي أدى إلى زيادة الارتباط بينهما .

2.4 معامل الارتباط (Coefficient of Correlation)

يعني وجود الارتباط بين ظاهرتين أن التغير " بالنقص أو الزيادة " في أحدهما يكون مصحوباً بتغير في الظاهرة الأخرى ، ويكون هذا التغير في نفس الاتجاه في حالة الارتباط الطردي ، وفي الاتجاه المخالف في حالة الارتباط العكسي ، أي أن الارتباط يمكن قياسه بواسطة التغيرات التي تحدث في الظاهرتين .

فإذا كان لدينا المتغيرين x ، y يعبران عن ظاهرتين معينتين فإن أفضل طريقة لمقارنة التغير في هاتين الظاهرتين هي مقارنة القيم المعيارية لهما أي القيم :

$$\left(\frac{x - \bar{x}}{\delta_x} \right), \left(\frac{y - \bar{y}}{\delta_y} \right)$$

حيث أن :

S_x ، S_y - هما الانحرافان المعياريان لقيم x ، y على الترتيب .

وهنا نلاحظ أن حاصل ضرب القيم المعيارية للظاهرتين يكون كبيراً عددياً " بغض النظر عن الإشارة موجبة كانت أم سالبة " في حالة وجود ارتباط قوي بين الظاهرتين وعليه فقد أتفق على اتخاذ متوسط حاصل ضرب

القيم المعيارية كمقياس لدرجة الارتباط بين المتغيرين والذي يسمى بمعامل الارتباط (Coefficient of Correlation) ويرمز له بالرمز R حيث أن :

$$R = \frac{1}{n} \sum \left(\frac{x - \bar{x}}{S_x} \right) \left(\frac{y - \bar{y}}{S_y} \right) \dots\dots\dots(1-4)$$

حيث أن :

n - هي عدد أزواج المفردات .

أن معامل الارتباط (R) يتميز بالخصائص التالية :

- 1- تتراوح قيمته العددية بين الصفر والواحد الصحيح .
- 2- هذا المقياس يساوي صفراً في حالة انعدام الارتباط ويساوي الواحد في حالة الارتباط التام .
- 3- تكون قيمة هذا المقياس موجبة حينما يكون الارتباط طردي وتكون سالبة في حالة الارتباط العكسي .
- 4- قيمة هذا المقياس العددية تزداد كلما ازدادت درجة الارتباط .

3.4 حساب معامل الارتباط (Calculation of Correlation Factor)

تعرف العلاقة السابقة (1-4) بمعامل بيرسون للارتباط (Pearson's Correlation) ويمكن كتابتها على الشكل التالي :

$$\therefore R = \frac{1}{n} \frac{\sum (x - \bar{x})(y - \bar{y})}{S_x \cdot S_y} \dots\dots\dots(2-4)$$

حيث أن كلاً من S_x ، S_y مقدار ثابت ويمكن أخذه كعامل مشترك في المقام .

وهذه العلاقة وإن كانت أسهل في حسابها من العلاقة السابقة إلا أنها تتطلب الكثير من العمليات الحسابية ، وخاصة إذا احتوى كل من \bar{x} و \bar{y} على كسور وما يترتب على ذلك من صعوبات في العمليات الحسابية ولغرض تبسيط الصيغة الأخيرة نقوم باختزال العلاقة (4 - 2) إلى العلاقة التالية :

$$\begin{aligned}
 R &= \frac{1}{n} \frac{\sum (x - \bar{x})(y - \bar{y})}{S_x \cdot S_y} = \frac{1}{n} \frac{\sum (x \cdot y - x \cdot \bar{y} - \bar{x} \cdot y + \bar{x} \cdot \bar{y})}{S_x \cdot S_y} \\
 R &= \frac{1}{n} \left(\frac{\sum x \cdot y - \sum x \cdot \bar{y} - \sum \bar{x} \cdot y + \sum \bar{x} \cdot \bar{y}}{S_x \cdot S_y} \right) \\
 &= \frac{\frac{\sum x \cdot y}{n} - \bar{y} \frac{\sum x}{n} - \bar{x} \frac{\sum y}{n} + \frac{n \bar{x} \cdot \bar{y}}{n}}{S_x \cdot S_y} \\
 &= \frac{\frac{\sum x \cdot y}{n} - \bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y} + \bar{x} \cdot \bar{y}}{S_x \cdot S_y} = \frac{\frac{1}{n} \sum x \cdot y - \bar{x} \cdot \bar{y}}{S_x \cdot S_x} \\
 \therefore R &= \frac{\frac{\sum x \cdot y}{n} - \bar{x} \cdot \bar{y}}{S_x \cdot S_y} \dots\dots\dots(3-4)
 \end{aligned}$$

حيث أن :

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{y} = \frac{\sum y}{n}$$

وبما أن :

$$\delta_x = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

وكذلك :

$$\delta_y = \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}$$

فأننا نحصل من خلال التعويض في المعادلة (3-4) على العلاقة التي تمكننا من حساب معامل الارتباط " معامل بيرسون " بطريقة سهلة حيث :

$$R = \frac{\frac{\sum x.y}{n} - \bar{x}.\bar{y}}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}} \dots\dots\dots(4-4)$$

مثال (1-4)

أحسب معامل ارتباط بيرسون بين قيم المتغيرين x و y من البيانات المبينة في الجدول (1-4) :

جدول (1-4)

65	68	62	70	66	67	64	68	71	69	x
28	29	26	28	25	28	25	31	30	28	y

الحل :

نقوم لحساب معامل الارتباط باستخدام العلاقة (4-4) ، حيث يلزمنا معرفة كل من المجاميع الآتية :

$$\sum x y \text{ و } \sum y^2 , \sum x^2 , \sum y , \sum x$$

وهذه يمكن حسابها من خلال الجدول (4-2) .

جدول (4-2)

إيجاد معامل الارتباط بين المتغيرين x و y

x	y	x^2	y^2	xy
69	28	4761	784	1932
71	30	5041	900	2130
68	31	4624	961	2108
64	25	4096	625	1600
67	28	4489	784	1876
66	25	4356	625	1650
70	28	4900	784	1960
62	26	3844	676	1612
68	29	4624	841	1972
65	28	4225	784	1820
$\sum x = 670$	$\sum y = 278$	$\sum x^2 = 44960$	$\sum y^2 = 7764$	$\sum xy = 18660$

$$\therefore \bar{x} = \frac{670}{10} = 67 \quad \bar{y} = \frac{287}{10} = 27.8$$

$$S_x = \sqrt{\frac{44960}{10} - \left(\frac{670}{10}\right)^2} = \sqrt{4496 - 4489} = \sqrt{7}$$

$$S_y = \sqrt{\frac{7764}{10} - \left(\frac{278}{10}\right)^2} = \sqrt{776.40 - 772.84} = \sqrt{3.56}$$

$$\therefore R = \frac{\frac{18660}{10} - (67)(27.8)}{(\sqrt{7})(\sqrt{3.56})} = \frac{1866 - 1862.6}{\sqrt{24.92}} = \frac{3.4}{4.99} = 0.68$$

ويطلق على معامل ارتباط بيرسون أيضاً معامل الارتباط العزمي ، ويمكن لحسابه استخدام العلاقة التالية :

$$R = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \dots\dots\dots(5-4)$$

ويشترط في استخدام القانون أن تكون العلاقة بين x و y خطية مع الأخذ بعين الاعتبار الملاحظات التالية :

1- في حالة كون معامل الارتباط سالب فإن العلاقة بين المتغيرين تكون سالبة تامة . إذا زاد أحد المتغيرين نقص الثاني بمقدار ثابت .

2- في حالة كون معامل الارتباط موجب فإن العلاقة بين المتغيرين تكون موجبة تامة .

3- أما في حالة كون معامل الارتباط = صفر فإن العلاقة بين المتغيرين تكون معدومة ويجب أن يكون معامل الارتباط محصوراً بين (- 1 إلى + 1) .

مثال (4-2)

أحسب معامل الارتباط العزمي لبيرسون للمتغيرين X و Y والمبينة في الجدول (4-3) :

جدول (4-3)

5	6	7	8	9	X
5	4	3	2	1	Y

الحل:

باستخدام العلاقة (4-5) ننظم الجدول (4-3) كما هو مبين في الجدول (4-4) :

جدول (4-4)

Y^2	X^2	$X \cdot Y$	Y	X
1	81	9	1	9
4	64	16	2	8
9	49	21	3	7
16	36	24	4	6
25	25	25	5	5
$\sum Y^2 = 55$	$\sum X^2 = 255$	$\sum XY = 95$	$\sum Y = 15$	$\sum X = 35$

$$\therefore R = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

$$\therefore R = \frac{5(95) - (35)(15)}{\sqrt{[(5)(255) - (35)^2][(5)(55) - (15)^2]}} = -1$$

إن العلاقة بين المتغيرين سالبة وتامة .

مثال (4-3)

أحسب معامل الارتباط بين المتغيرين x, y الموضحين في الجدول (5-4) :

جدول (5-4)

5	4	3	2	1	X
2	3	1	1	2	Y

الحل:

نضع البيانات الموجودة في الجدول (5-4) بالشكل العمودي الموضح في الجدول (6-4) :

جدول (6-4)

Y^2	X^2	XY	Y	X
9	1	3	3	1
1	4	2	1	2
1	9	3	1	3
9	16	12	3	4
4	25	10	2	5
$\sum Y^2 = 24$	$\sum X^2 = 55$	$\sum XY = 30$	$\sum Y = 10$	$\sum X = 15$

$$\therefore R = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

$$\therefore R = \frac{5(30) - (15)(10)}{\sqrt{[(5)(55) - (15)^2][(5)(24) - (10)^2]}} = 0$$

وعلى هذا الأساس فإن العلاقة تكون معدومة بين المتغيرين .

4.4 إيجاد معامل الارتباط باستخدام الوسط الفرضي (Calculation of Correlation Factor Using Assumed Mean)

أن العلاقة التي تم استخدامها في حل المثال (4-1) تتطلب الكثير من العمليات الحسابية ، وأن الحل سيكون أسهل بكثير إذا استخدمنا وسطين فرضيين للقيم x و y فإذا حسبنا انحرافات (x) عن الوسط الفرضي (a) وانحرافات (y) عن الوسط الفرضي (b) سنجد أن :

$$d_x = (x - a)$$

$$d_y = (y - b)$$

وقد سبق أن بينا أن البسط في صيغة معامل الارتباط هو :

$$\begin{aligned} & \frac{1}{n} \sum (x - \bar{x})(y - \bar{y}) \\ &= \frac{1}{n} \sum [(x - a) - (\bar{x} - a)][(y - b) - (\bar{y} - b)] \\ &= \frac{1}{n} \sum (d_x - \bar{d}_x)(d_y - \bar{d}_y) \quad \dots\dots\dots(4-6) \end{aligned}$$

حيث أن :

$$\bar{d}_x = \bar{x} - a$$

$$\bar{d}_y = \bar{y} - b$$

وعلى الطالب مراجعة علاقة حساب المتوسط الحسابي للبيانات المبوبة باستخدام الوسط الفرضي التي تم شرحها بالتفصيل في الباب الثاني . والنتيجة الأخيرة لبسط معامل الارتباط يمكن كتابتها على النحو التالي :

$$\frac{\sum d_x \cdot d_y}{n} - \bar{d}_x \cdot \bar{d}_y$$

وبالتالي نحصل على العلاقة التي تمكنا من حساب معامل الارتباط باستخدام وسطين فرضيين لقيم x وقيم y وهي بالشكل التالي :

$$\therefore R = \frac{\frac{\sum d_x \cdot d_y}{n} - \sum \bar{d}_x \cdot \bar{d}_y}{\delta_x \cdot \delta_y} \dots\dots\dots(7-4)$$

حيث أن :

$$\frac{\sum d_x}{n} = \bar{d}_x$$

$$\frac{\sum d_y}{n} = \bar{d}_y$$

أما الانحراف المعياري فيساوي :

$$S_x = \sqrt{\frac{\sum d_x^2}{n} - \left(\frac{\sum d_x}{n}\right)^2} \dots\dots\dots(8-4)$$

$$S_y = \sqrt{\frac{\sum d_y^2}{n} - \left(\frac{\sum d_y}{n}\right)^2} \dots\dots\dots(9-4)$$

ويمكن حساب معامل الارتباط بين قيم x ، y في المثال السابق باستخدام هذه العلاقة . وذلك باعتبار القيمة 68 كوسط فرضي لقيم x والقيمة 28 كوسط فرضي لقيم y ، وقد اخترنا هاتين القيمتين نظراً لتكرارهما مما يسهل الحل كما هو مبين في الجدول (7-4) .

جدول (7-4)

إيجاد معامل الارتباط بين قيم x ، y باستخدام وسطين فرضيين

$d_x d_y$	d_y^2	d_x^2	$d_y (y - 28)$	$d_x (x - 68)$	y	x
0	0	1	0	1	28	69
6	4	9	2	3	30	71
0	9	0	3	0	31	68
12	9	16	3 -	4 -	25	64
0	0	1	0	1 -	28	67
6	9	4	3 -	2 -	25	66
0	0	4	0	2	28	70
12	4	36	2 -	6 -	26	62
0	1	0	1	0	29	68
0	0	9	0	3 -	28	65
36	36	80	2 -	10 -	المجموع	

$$\therefore \bar{d}_x = \frac{\sum d_x}{n} = \frac{10 -}{10} = -1$$

$$\therefore \bar{d}_y = \frac{\sum d_y}{n} = \frac{2 -}{10} = -0.2$$

$$\therefore S_x = \sqrt{\frac{\sum d_x^2}{n} - \left(\frac{\sum d_x}{n}\right)^2} = \sqrt{\frac{8}{10} - \left(\frac{10}{10}\right)^2}$$

$$= \sqrt{8 - 1} = \sqrt{7}$$

$$S_y = \sqrt{\frac{\sum d_y^2}{n} - \left(\frac{\sum d_y}{n}\right)^2} = \sqrt{\frac{36}{10} - \left(\frac{2}{10}\right)^2}$$

$$= \sqrt{3.60 - 0.04} = \sqrt{3.56}$$

$$\therefore R = \frac{\frac{\sum d_x \cdot d_y}{n} - \sum \bar{d}_x \cdot \bar{d}_y}{S_x \cdot S_y} = \frac{\frac{36}{10} - (-1)(-0.2)}{(\sqrt{7})(\sqrt{3.56})}$$

$$= \frac{3.6 - 0.2}{\sqrt{24.92}} = \frac{3.4}{4.99} = 0.68$$

وهي نفس النتيجة التي حصلنا عليها من قبل مع ملاحظة قلة العمليات الحسابية ، ولذلك يجب استخدام هذه العلاقة الأخيرة لتسهيل العمل الحسابي ، فإذا طلب منا حساب معامل بيرسون للارتباط فيجب استخدام هذه الصيغة نظراً لسهولة حسابها .

مثال (4-4)

أحسب معامل ارتباط بيرسون بين قيم x ، y من البيانات المبينة في الجدول (8-4) :

جدول (8-4)

164	154	170	169	170	170	165	x
62	56	83	65	72	70	61	y

الحل:

باستخدام القيمة 170 كوسط فرضي لقيم x والقيمة 65 كوسط فرضي لقيم y ومن الجدول (8-4) يمكن تكوين جدول (9-4) الخاص بالمثال على الشكل التالي :

جدول (9-4)

$d_x d_y$	d_y^2	d_x^2	$d_y (y - 28)$	$d_x (x - 68)$	y	x
20	16	25	4 -	5 -	61	165
0	25	0	5	0	70	170
0	49	0	7	0	72	170
0	0	1	0	1 -	65	169
0	324	0	18	0	83	170
144	81	256	9 -	16 -	56	154
18	9	36	3 -	6 -	62	164
182	504	318	14	28 -	المجموع	

$$\therefore \bar{d}_x = \frac{\sum d_x}{n} = \frac{-28}{7} = -4$$

$$\therefore \bar{d}_y = \frac{\sum d_y}{n} = \frac{14}{7} = 2$$

$$\begin{aligned} \therefore S_x &= \sqrt{\frac{\sum d_x^2}{n} - \left(\frac{\sum d_x}{n} \right)^2} = \sqrt{\frac{318}{7} - \left(\frac{28 -}{7} \right)^2} \\ &= \sqrt{45.43 - 16} = \sqrt{29.43} \end{aligned}$$

$$\begin{aligned} S_y &= \sqrt{\frac{\sum d_y^2}{n} - \left(\frac{\sum d_y}{n} \right)^2} = \sqrt{\frac{504}{7} - \left(\frac{14}{7} \right)^2} \\ &= \sqrt{72 - 4} = \sqrt{68} \end{aligned}$$

$$\therefore R = \frac{\frac{\sum d_x \cdot d_y}{n} - \bar{d}_x \cdot \bar{d}_y}{\delta_x \cdot \delta_y} = \frac{\frac{182}{7} - (-4)(2)}{(\sqrt{29.43})(\sqrt{68})}$$

$$= \frac{26 + 8}{\sqrt{2001.24}} = \frac{34}{44.55} = 0.76$$

إن هناك ارتباط طردي قوي .

5.4 حساب معامل الارتباط للبيانات المبوبة

(Correlation Factor for Tabulated Data)

عندما يكبر حجم العينة يكون من الصعب حساب معامل الارتباط بالطريقة السابقة ، لذلك يمكن تبويب البيانات في شكل جدول توزيع تكراري مزدوج والذي تم عرضه في الباب الأول من هذا الكتاب . ومن ثم يتم إيجاد معامل الارتباط باستخدام العلاقة التالية :

$$R = \frac{\frac{\sum (d_x \cdot d_y) \cdot f}{n} - \bar{d}_x \cdot \bar{d}_y}{S_x \cdot S_y} \dots\dots\dots(10-4)$$

حيث أن :

$$n = \sum f_i$$

$$\bar{d}_x = \frac{\sum d_x \cdot f}{\sum f}$$

$$\bar{d}_y = \frac{\sum d_y \cdot f}{\sum f}$$

أما :

$$S_x = \sqrt{\frac{(\sum d_x^2) f_i}{\sum f} - \left(\frac{\sum d_x \cdot f_i}{\sum f} \right)^2} \dots\dots\dots(11-4)$$

$$S_y = \sqrt{\frac{(\sum d_y^2) f_i}{\sum f} - \left(\frac{\sum d_y \cdot f_i}{\sum f} \right)^2} \dots\dots\dots(12-4)$$

حيث أن :

d_x ، d_y - هي انحرافات مراكز الفئات للمتغيرين x و y عن وسطيهما
الفرضيين .

وفي حالة تساوي أطوال فئات كل من المتغيرين فإن استخدام الانحرافات
المختصرة بدلاً من الانحرافات لن يؤثر على النتيجة وبذلك يمكن استعمال
العلاقة التالية :

$$R = \frac{\frac{\sum (\bar{d}_x \cdot \bar{d}_y) \cdot f}{\sum f} - \left(\frac{\sum \bar{d}_x \cdot f_x}{\sum f_x} \right) \left(\frac{\sum \bar{d}_y \cdot f_y}{\sum f_y} \right)}{\sqrt{\frac{\sum (\bar{d}_x^2 \cdot f_x)}{\sum f_x} - \left(\frac{\sum \bar{d}_x \cdot f_x}{\sum f_x} \right)^2} \sqrt{\frac{\sum (\bar{d}_y^2 \cdot f_y)}{\sum f_y} - \left(\frac{\sum \bar{d}_y \cdot f_y}{\sum f_y} \right)^2}} \dots\dots\dots(13-4)$$

مثال (5-4)

جدول (10-4) التكراري المزدوج يمثل العلاقة بين الطول والوزن لعينة
مكونة من 100 طالب من إحدى المدارس . أحسب معامل ارتباط بيرسون
لذلك البيانات .

جدول (10-4)

المجموع	175 - 170	- 165	- 160	- 155	- 150	الطول الوزن
8				3	5	- 40
27			14	12	1	- 50
52		22	28	2		- 60
13	2	8	3			80 - 70
100	2	30	45	17	6	المجموع

الحل:

نلاحظ هنا تساوي أطوال الفئات بالنسبة للطول (5cm) والوزن (10kg) وبذلك يمكننا استخدام الصيغة السابقة . ولحسابها نضيف إلى الجدول سبعة صفوف وسبعة أعمدة كما في الجدول (4-11) . ففي العمود الأول نحسب مراكز فئات المتغير (y) ، ثم نختار من بينها وسطاً فرضياً ونحسب الانحرافات عنه d_y ، وفي العمود التالي وحيث أن أطوال الفئات متساوية نقسم كل من الانحرافات d_y على طول الفئة فنحصل على \bar{d}_y في العمود الثالث .

ونلاحظ أنها تساوي صفر أمام الفئة التي اختير مركزها كوسط فرضي -1 ، - 2 في الفئات السابقة لها ، 1 ، 2 ، 3 في الفئات اللاحقة لها ، وفي العمود الرابع نضرب كل انحراف مختصر في التكرار المناظر له فنحصل على $\bar{d}_y \cdot f_i$ وفي العمود الخامس نضرب كل انحراف مختصر \bar{d}_y (العمود الثالث) في $\bar{d}_y \cdot f_y$ (العمود الرابع) فنحصل على

$\bar{d}_Y^2 \cdot f_Y$ وبتكرار نفس الخطوات بالنسبة للصفوف نحصل على مراكز فئات المتغير x وهي :

$$\bar{d}_X^2 \cdot f_X, \bar{d}_X \cdot f_X, \bar{d}_X, d_X$$

في الخمسة صفوف التي أضفناها ، والبيانات التي حسبناها حتى الآن تكفي لحساب المقام في صيغة معامل الارتباط .

وقبل حساب بيانات العموديين السادس والسابع ننقل بيانات الانحرافات المختصرة \bar{d}_Y خارج الجدول الأصلي (إلى يمين الفئات y) والانحرافات المختصرة \bar{d}_X خارج الجدول (فوق الفئات x) ونحسب قيم العمود السادس بأن نضرب التكرارات في كل صف (من الجدول الأصلي) في الانحرافات المناظرة والتي كتبناها أعلى للجدول ثم نجمع النتائج . فأول قيمة في العمود السادس تكون :

$$13 - = (1-)(3) + (2-)(5)$$

وثاني قيمة تكون :

$$14 - = (14)(صفر) + (1-)(12) + (2-)(1)$$

وبذلك نحصل على بيانات العمود السادس التي تعطي لنا $\sum \bar{d}_Y \cdot f_Y$ أما بيانات العمود السابع فنحصل عليها بضرب كل قيمة من قيم العمود السادس في الانحراف المناظر، وهو \bar{d}_X فنحصل على $\sum \bar{d}_X \cdot \bar{d}_Y \cdot f$ وبذلك يمكننا حساب معامل الارتباط الآن . أما الصفين السادس والسابع فيمكن تكملتهما الحل بدونهما ولكن الهدف منهما هو التأكد من صحة العمليات الحسابية السابقة . فالصف السادس نحسبه بنفس الطريقة

التي حسبت بها بيانات العمود السادس ، ونحصل على قيمته. بأن نضرب التكرارات في كل عمود من الجدول الأصلي في الانحرافات المناظرة والتي كتبناها إلى يمين الجدول ثم نجمع النتائج . فأول قيمة في الصف السادس تكون :

$$11-(1-)(1)+(2-)(5)$$

وثاني قيمة فيه تكون:

$$18-(3-)(2)+(1-)(12)+(1-)(12)+(2-)(12)$$

وبذلك نحصل على بيانات الصف السادس وتعطي لنا :

$$\sum \bar{d}_X \cdot f_X$$

أما بيانات الصف السابع فنحصل عليها بضرب كل قيمة من قيم الصف السادس في الانحراف المناظر \bar{d}_Y ونكون بذلك قد حصلنا على :

$$\sum \bar{d}_X \cdot \bar{d}_Y \cdot f$$

نلاحظ هنا أن مجموع الصف الرابع لا بد وأن يساوي مجموع العمود السادس وأن مجموع الصف السادس لا بد وأن يساوي مجموع العمود الرابع وأن مجموع الصف السابع سوف يساوي مجموع العمود السابع كما يتضح من الأسهم في الجدول (4-11) ، وبعد تكملة الجدول بالشكل الذي شرحناه يمكن حساب معامل الارتباط :

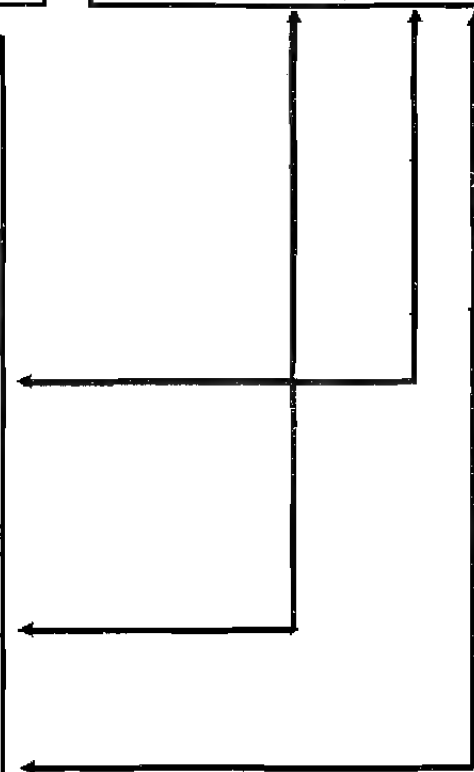
جدول (4-11) إيجاد معامل الارتباط من الحدود التكراري المزدوج

2 - 1 - 0 1 - 2 -

$\sum \bar{d}_x \cdot \bar{d}_y \cdot f$	$\sum \bar{d}_x \cdot f$	$\bar{d}_x^2 \cdot f_x$	$\bar{d}_y \cdot f_y$	d_x	d_y	مركز الفئة y
26	13 -	32	16 -	2 -	20 -	55
14	14 -	27	27	1	10 -	68
0	20	0	0	0	0	78
12	12	13	13	1	10	85
52	5	72	30 -			---

f_y	175 - 170	- 165	160	- 155	- 150	X / Y
8				3	5	- 40
27			14	12	1	- 50
52		22	28	2		- 60
13	2	8	3			80 - 70
100	2	30	45	17	6	f_x

	172.5	167.5	162.5	157.5	152.5	مركز الفئة x
----	10	5	0	5 -	10 -	d_x
----	2	1	0	1 -	2 -	\bar{d}_x
5	4	30	0	17 -	12 -	$\bar{d}_x \cdot f_x$
79	8	30	0	17	24	$\bar{d}_x^2 \cdot f_x$
30 -	2	8	11 -	18 -	11 -	$\sum \bar{d}_y \cdot f$
52		8	0	18	22	$\sum \bar{d}_x \cdot \bar{d}_y \cdot f$



$$\begin{aligned}
\therefore R &= \frac{\frac{52}{100} - \left(\frac{5}{100}\right)\left(\frac{30-}{100}\right)}{\sqrt{\frac{79}{100} - \left(\frac{5}{100}\right)^2} \sqrt{\frac{72}{100} - \left(\frac{30-}{100}\right)^2}} \\
&= \frac{0.52 + (0.05)(0.3)}{(\sqrt{0.7875})(\sqrt{0.63})} \\
&= \frac{0.52 + 0.15}{\sqrt{0.496125}} \\
&= \frac{0.535}{0.7043} = 0.76
\end{aligned}$$

إن يوجد ارتباط طردي قوي بين الطول والوزن لعينة الطلبة المدروسة . وتجد الإشارة إلى أنه يمكن اختصار الحل في حالة تساوي أطوال الفئات وذلك بعدم إضافة كل من العموديين الأول والثاني والصفين الأول والثاني في الجدول (4-11) ، وذلك بأن نكتب الانحرافات المختصرة مباشرة بوضع صفر أمام الفئة التي كنا سنختار مركزها كوسط فرضي (ويفضل أن تكون في منتصف الجدول وأمام أكبر تكرار) ثم نكتب -1 ، -2 ... الانحرافات المختصرة للفئات السابقة لها، 1 ، 2 ، 3 ... للانحرافات المختصرة للفئات اللاحقة لها ثم نكمل الحل .

مثال (4-6)

أحسب معامل الارتباط بين قيم x ، y من البيانات المبينة في الجدول (4-12) .

الجدول (12-4)

المجموع	90 - 80	- 70	- 60	- 50	- 40	Y / X
6	—	—	—	—	6	- 40
22	—	—	4	12	6	- 50
42	—	2	20	10	—	- 60
30	2	16	12	—	—	80 - 70
100	2	18	46	22	12	المجموع

الحل :

نضيف إلى الجدول خمسة صفوف وخمسة أعمدة ونتبع نفس خطوات المثال السابق فنجد أن :

$$\sum \bar{d}_X \cdot \bar{d}_Y \cdot f = 68 \quad , \quad \sum \bar{d}_X \cdot f_X = -24 \quad , \quad \sum \bar{d}_Y \cdot f_Y = -4$$

وأن :

$$\sum \bar{d}_Y^2 \cdot f_Y = 76 \quad , \quad \sum \bar{d}_X^2 \cdot f_X = 96$$

كما مبين في الجدول (4 - 13) .

$$\begin{aligned} \therefore R &= \frac{\frac{68}{100} - \left(\frac{24}{100} \right) \left(\frac{4}{100} \right)}{\sqrt{\frac{96}{100} - \left(\frac{24}{100} \right)^2} \sqrt{\frac{76}{100} - \left(\frac{4}{100} \right)^2}} \\ &= \frac{0.68 - (0.24)(0.04)}{(\sqrt{0.9024})(\sqrt{0.7584})} \\ &= \frac{0.68 - 0.0096}{\sqrt{0.6838}} \\ &= \frac{0.6704}{0.8272} = 0.81 \end{aligned}$$

جدول (4 - 13) إيجاد معامل الارتباط من الجدول التكراري المزدوج

		2		1		0		1		2		X \ Y
		2		1		0		1		2		
		90 - 80		- 70		- 60		- 50		- 40		Y
f_Y												
6										6		- 40
22						4		12		6		- 50
42				2		30		10				- 60
30		2		16		12						80 - 70
100		2		18		46		22		12		f_X

$\sum \bar{d}_x \cdot \bar{d}_y \cdot f$	$\sum \bar{d}_x \cdot f$	$\bar{d}_x^2 \cdot f_Y$	$\bar{d}_x \cdot f_Y$	\bar{d}_y
24	12 -	24	12 -	2 -
24	24 -	22	22 -	1 -
0	8 -	0	0	0
20	20	30	30	1
68	24 -	76	4 -	

	2	1	0	1	2
\bar{d}_x	2	1	0	1	2
$\bar{d}_x \cdot f_Y$	24 -	18	0	22 -	12 -
$\bar{d}_x^2 \cdot f_Y$	96	18	0	22	24 -
$\sum \bar{d}_x \cdot f$	4 -	16	8	12 -	18 -
$\sum \bar{d}_x \cdot \bar{d}_y \cdot f$	68	16	0	12	36

إنه هناك ارتباط طردي قوي بين قيم x و y .

6.4 معامل ارتباط الرتب (سبيرمان)

(Spearman's Order Correlation Factor)

يستخدم هذا المعامل لدراسة الارتباط بين البيانات النوعية ، أي تلك التي لا يمكن قياسها كمياً . وتعتمد هذه الطريقة على إعطاء المتغيرات رتباً لتحل محل القياس العددي . فإذا رتبنا مفردات المتغير x ترتيباً تصاعدياً ووجدنا أن مفردات المتغير y المناظرة لها مرتبة ترتيباً تصاعدياً أيضاً نستنتج وجود ارتباط طردي تام بين المتغيرين x ، y . أما إذا رتبنا مفردات المتغير x ترتيباً تصاعدياً ووجدنا أن مفردات المتغير y المناظرة لها مرتبة ترتيباً تنازلياً نستنتج وجود ارتباط عكسي تام بين المتغيرين x ، y غير أن هذا الارتباط التام نادراً ما يصادفنا في الدراسات الاجتماعية والاقتصادية .

ولقياس معامل الارتباط بين مفردات المتغيرين x و y نرتب كلا منهما حسب أفضليته ثم نحسب الفرق بين كل رتبتين متتاليتين فنجد أن :

$$\sum F = 0$$

وبحساب مربعات هذه الفروقات يمكن إيجاد معامل الارتباط باستخدام العلاقة :

$$R = 1 - \frac{6 \sum F^2}{n(n^2 - 1)} \dots\dots\dots (14 - 4)$$

حيث أن :

n = عدد الرتب .

$\sum F$ = مجموع مربعات الفروقات بين الرتب .

مثال (7-4)

الجدول (14-4) يبين تقديرات ستة من الطلبة في امتحان مادتي الفيزياء والكيمياء أحسب معامل ارتباط الرتب لسبيرمان بين تقديرات المادتين .

جدول (14-4)

رقم الطالب	1	2	3	4	5	6
تقدير الفيزياء	ضعيف	ممتاز	جيد	ضعيف جداً	مقبول	جيد جداً
تقدير الكيمياء	مقبول	جيد جداً	جيد	ضعيف	ضعيف جداً	ممتاز

الحل:

لحساب معامل الارتباط من هذه البيانات نرتب تقديرات كل من المادتين ترتيباً تصاعدياً أو تنازلياً وذلك بإعطاء التقدير ممتاز الرتبة (1) والتقدير الذي يليه الرتبة (2) و..... هكذا ثم نحسب الفروقات بين كل رتبتين متناظرتين كما في جدول (15-4) فنجد أن :

$$\begin{aligned} \therefore R &= 1 - \frac{6 \sum F^2}{n(n^2 - 1)} = 1 - \frac{(6)(8)}{(6)(36 - 1)} = 1 - \frac{8}{35} \\ &= \frac{35 - 8}{35} = \frac{27}{35} = 0.77 \end{aligned}$$

إذن يوجد ارتباط طردي قوي بين تقديرات الطلبة الستة في هاتين المادتين .

جدول (15-4)

حساب معامل ارتباط الرتب - سبيرمان - بين تقديرات مادتي الفيزياء والكيمياء

تقدير الفيزياء	تقدير الكيمياء	رتب تقدير الفيزياء	رتب تقدير الكيمياء	الفروق F	مربع الفروق F ²
ضعيف	مقبول	5	4	1	1
ممتاز	جيد جداً	1	2	1-	1
جيد	جيد	3	3	صفر	صفر
ضعيف جداً	ضعيف	6	5	1	1
مقبول	ضعيف جداً	4	6	2-	4
جيد جداً	ممتاز	2	1	1	1
المجموع					8

7.4 معامل ارتباط الرتب (سبيرمان) في حالة الرتب التكرارية

(Spearman's Correlation Coefficient for Frequent Orders)

في المثال السابق نلاحظ أنه لم تتكرر أي من التقديرات التي حصل عليها الطلبة . فإذا صانفنا مثلاً آخرًا تتكرر فيه بعض التقديرات فإننا نعطي القيم المتكررة رتباً تساوي متوسط الرتب التي كانت ستعطي لو لم تتكرر التقديرات .

مثال (8-4)

الجدول (16-4) يبين تقديرات عشرة من الطلبة في امتحان مادتي الإحصاء والاقتصاد والمطلوب حساب معامل الارتباط بين تقديرات المادتين .

جدول (4-16)

رقم الطالب	1	2	3	4	5	6	7	8	9	10
تقدير الإحصاء	ضعيف جداً	مقبول	ممتاز	مقبول	ضعيف	جيد جداً	جيد	جيد	مقبول	مقبول
تقدير الاقتصاد	مقبول	جيد	جيد جداً	مقبول	جيد	مقبول	ممتاز	ضعيف جداً	ضعيف	جيد جداً

الحل :

عند إعطاء رتب التقدير لمادة الإحصاء سنجد أن الطالب رقم 3 يأخذ الرتبة (1) والطالب رقم 6 يأخذ الرتبة (2) والطالب رقم 7 يأخذ الرتبة (3) بينما الطلبة رقم 2 ، 4 ، 9 ، 10 لهم نفس التقدير ويستحقون الرتب (4) ، (5) ، (6) ، (7) ونظراً لتقاربهم في التقدير يعطى لكل منهم متوسط هذه الرتب وهو حاصل قسمة مجموع هذه الرتب على عددها أي:

$$5.5 = \frac{7+6+5+4}{4}$$

ويلي ذلك الطالبان رقم (5) ، (8) ولما كان لكل منهما نفس التقدير لذلك يعطى لكل منهما متوسط الرتبتين أي :

$$8.5 = \frac{9+8}{2}$$

ويلي ذلك الطالب رقم 1 حيث يأخذ الرتبة (10) ، وبإتباع نفس الطريقة عند إعطاء رتب التقدير لمادة الاقتصاد يمكن أن نحسب الفروقات كما موضح في الجدول (4-17) ثم نكمل حل المسألة بالطريقة المعتادة فنجد أن :

جدول (4-17)

حساب معامل الارتباط لسبيرمان في حالة الرتب التكرارية

رقم الطالب	تقدير الإحصاء	تقدير الاقتصاد	رتب تقديري الإحصاء	رتب تقديري الاقتصاد	الفرق F	مربع الفرقات F ²
1	ضعيف جداً	مقبول	10	7	3	9
2	مقبول	جيد	5.5	4.5	1	1
3	ممتاز	جيد جداً	1	2.5	-1.5	2.25
4	مقبول	متبرل	5.5	7	-1.5	2.25
5	ضعيف	جيد	8.5	4.5	4	16
6	جيد جداً	مقبول	2	7	-5	25
7	جيد	ممتاز	3	1	2	4
8	ضعيف	ضعيف جداً	8.5	10	-1.5	2.25
9	مقبول	ضعيف	5.5	9	-3.5	12.25
10	مقبول	جيد جداً	5.5	2.5	3	9
المجموع						83.00

$$\therefore R = 1 - \frac{6 \sum F^2}{n(n^2 - 1)} = 1 - \frac{(6)(83)}{(10)(100 - 1)} = 1 - \frac{(6)(83)}{(10)(99)}$$

$$= 1 - \frac{498}{990} = 1 - 0.503 = 0.497$$

وهذه القيمة لمعامل الارتباط تبين أن هناك ارتباطاً طردياً ليس بالقوي وليس بالضعيف .

ويجب ملاحظة أنه لا يقتصر استخدام معامل سبيرمان للارتباط على المتغيرات غير القابلة للقياس الكمي (كما تم توضيحه من خلال حل المثالين السابقين ولكن قد يستخدم أيضاً لحساب الارتباط بين المتغيرات القابلة للقياس الكمي وذلك رغبة في تقليل واختصار العمليات الحسابية كما يتضح من حل المثال (9-4) .

مثال (9-4)

أحسب معامل الارتباط لسبيرمان بين قيم x , y من البيانات المبينة في الجدول (18-4) :

جدول (18-4)

15	14	12	14	11	x
18	13	14	13	12	y

الحل :

من الجدول (18-4) نعطي المتغيرين x , y رتباً ثم نحسب الفروقات بين الرتب المتقابلة ونوجد مربعاتها كما هو مبين في الجدول (19-4) .

جدول (19-4)

F^2	F	رتب y	رتب x	y	x
0	0	5	5	12	11
1	1 -	3.5	2.5	13	14
4	2	2	4	14	13
1	1 -	3.5	2.5	13	14
0	0	1	1	18	15
6	المجموع				

$$\therefore R = 1 - \frac{6 \sum F^2}{n(n^2 - 1)} = 1 - \frac{(6)(6)}{(5)(25 - 1)} = 1 - \frac{(6)(6)}{(5)(24)}$$

$$= 1 - \frac{3}{10} = 1 - 0.3 = 0.7$$

إن وجود ارتباط طردي قوي بين y ، x

مثال (4- 10)

إذا كانت تقديرات 5 طلبة في مادتين دراسيتين مبينة في الجدول (20-4) المطلوب حساب معامل الارتباط بين هاتين المادتين .

جدول (20-4)

الطالب	1	2	3	4	5
تقدير المادة الأولى	جيد	ممتاز	جيد جداً	مقبول	ضعيف
تقدير المادة الثانية	مقبول	جيد جداً	جيد	ممتاز	ضعيف

الحل:

لاحظ أن البيانات وصفية خاضعة للترتيب وبالتالي فإن معامل الارتباط المناسب لها هو معامل ارتباط الرتب لسبيرمان وبحسب كما هو مبين في الجدول (21-4) :

جدول (21-4)

الطالب	تقدير العادة الأولى	تقدير العادة الثانية	رتب تقدير العادة الثانية	رتب تقدير العادة الأولى	الفرق F	مربعات الفرق F^2
1	جيد	مقبول	4	3	1 -	1
2	ممتاز	جيد جداً	2	1	1 -	1
3	جيد جداً	جيد	3	2	1 -	1
4	مقبول	ممتاز	1	4	3	9
5	ضعيف	ضعيف	5	5	0	0
المجموع			15	15	0	12

$$\therefore R = 1 - \frac{6 \sum F^2}{n(n^2 - 1)} = 1 - \frac{(6)(12)}{(5)(25 - 1)} = 1 - \frac{(6)(12)}{(5)(24)}$$

$$= 1 - \frac{3}{5} = 1 - 0.6 = 0.4$$

إنه يوجد ارتباط طردي ولكنه ليس قوياً .

مثال (11-4)

ذكر أحد المختصين بأن الارتباط بين عدد الأطفال والمستوى التعليمي لرب الأسرة ارتباط طردي قوي ، فهل تؤيد رأيه بناءً على البيانات التالية والمبينة في الجدول (22-4).

جدول (4-22)

تطالب	تقدير المادة الأولى	تقدير المادة الثانية	ركب تقدير المادة الثانية	ركب تقدير المادة الأولى	الفرق F	مربعات الفرق F^2
1	جيد	مقبول	4	3	1 -	1
2	ممتاز	جيد جداً	2	1	1 -	1
3	جيد جداً	جيد	3	2	1 -	1
4	مقبول	ممتاز	1	4	3	9
5	ضعيف	ضعيف	5	5	0	0
المجموع		15	15	0	12	

الأسرة	1	2	3	4	5	6	7	8	9
عدد الأطفال	4	3	2	7	3	4	7	3	5
المستوى التعليمي	يقراً ويكتب	شهادة عليا	شهادة متوسطة	شهادة متوسطة ويقرأ ويكتب	شهادة متوسطة	شهادة متوسطة	أمي	شهادة متوسطة	أمي

الحل:

نلاحظ أنه لا يمكن حساب معامل ارتباط الرتب لسبب عدم وجود بيانات إلا بعد تعديل الرتب للصفات المكررة أي الحصول على الرتب المعدلة وذلك باتباع الآتي كما هو مبين في الجدول (4-23) .

جدول (4-23)

عدد الظلال	المستوى الظلي	رتب عدد الظلال	رتب المستوى الظلي	الرتب المعطاة لعدد الظلال	الرتب المعطاة للمستوى الظلي	الفرق F	مربعات الفرق
4	بقرأ ويكتب	5	3	5.5	3.5	2	4
3	شهادة عليا	2	9	3	9	-6	36
2	شهادة متوسطة	1	5	1	6.5	-5.5	30.25
7	بقرأ ويكتب	8	4	8.5	3.5	5	25
3	شهادة متوسطة	3	6	3	6.5	-3.5	12.25
4	شهادة متوسطة	6	7	5.5	6.5	-1	1
7	أسي	9	1	8.5	1.5	7	49
3	شهادة متوسطة	4	8	3	6.5	-3.5	12.25
5	أسي	7	2	7	1.5	5.5	30.25
المجموع		45	45	45	45	0	200

$$\therefore R = 1 - \frac{6 \sum F^2}{n(n^2 - 1)} = 1 - \frac{(6)(200)}{(9)(81 - 1)} = 1 - \frac{(6)(200)}{(9)(80)}$$

$$= 1 - \frac{5}{3} = 1 - 1.67 = 0.67 -$$

وسنلاحظ أن الارتباط قوي إلى حد ما وعكسي وهذا لا يؤيد رأي الإحصائي الاجتماعي .

8.4 معامل الاقتران (Association Coefficient)

يستخدم معامل الاقتران لدراسة قوة واتجاه العلاقة بين ظاهرتين (لا يمكن التعبير عنهما رقمياً وغير خاضعتين لترتيب) ولكل منهما صفتان متعاكستان مثل موجود وغير موجود ، وعادة ما تظهر البيانات في هذه الحالة كما هو مبين في الجدول (4-24) .

جدول (24-4)

المجموع	الصفة الأولى	الصفة الثانية	الظاهرة الثانية
			الظاهرة الأولى
$k_1 = a_{11} + a_{12}$ $k_2 = a_{21} + a_{22}$	a_{12}	a_{11}	الصفة الأولى
	a_{22}	a_{21}	الصفة الثانية
$k_1 + k_2$	$L_2 = a_{12} + a_{22}$	$L_1 = a_{11} + a_{21}$	المجموع

يحسب معامل الاقتران (A) من القانون التالي:

$$A = \frac{a_{11} a_{22} - a_{12} a_{21}}{a_{11} a_{22} + a_{12} a_{21}} \dots\dots\dots(15-4)$$

مثال (12-4)

لتحديد قوة واتجاه العلاقة بين حالة التدخين والمستوى التعليمي للأشخاص ، جمعت بيانات من 100 شخص والمبينة في الجدول (25-4) على النحو التالي :

جدول (25-4)

المجموع	غير متعلم	متعلم	المستوى التعليمي
			حالة التدخين
50	20	30	مدخن
50	10	40	غير مدخن
100	30	70	المجموع

احسب معامل الاقتران بين المستوى التعليمي وظاهرة التدخين .

الحل:

حيث أن لدينا ظاهرتين فقط هما حالتى التدخين والمستوى التعليمي ولكل منهما صفتان متعاكستين لذا يستخدم معامل الاقتران على النحو التالي :

$$A = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11}a_{22} + a_{12}a_{21}} = \frac{(30)(10) - (40)(20)}{(30)(10) + (40)(20)} = \frac{300 - 800}{300 + 800} \\ = \frac{-500}{1100} = -0.46$$

وبالتالي فإن معامل الاقتران بين المستوى التعليمي وظاهرة التدخين اقتران عكسي ليس بالقوي . أي الاقتران بين المتعلم وغير المدخن أو الاقتران بين غير المتعلم والمدخن .

مثال (4-13)

سئل 60 رجلاً عن رأيهم في حق المرأة في العمل فأجاب 47 رجلاً منهم بالرفض . وسئلت 40 امرأة فأجابت 5 منهن بالرفض . أدرس قوة واتجاه العلاقة بين نوع الجنس وحق المرأة في العمل .

الحل:

حيث إنه لدينا ظاهرتين وصفيتين لكل منهما صفتان متعاكستان وبالتالي نستخدم معامل الاقتران وكما هو مبين في الجدول (4-26).

جدول (4-26)

المجموع	أنثى	ذكر	حق المرأة في العمل ↓ الجنس ←
			مؤيد رافض
48	35	13	
52	5	47	
100	40	60	المجموع

يحسب معامل الاقتران من المعادلة (4-27) حيث :

$$\begin{aligned}
 A &= \frac{a_{11} a_{22} - a_{12} a_{21}}{a_{11} a_{22} + a_{12} a_{21}} \\
 &= \frac{(13)(5) - (35)(47)}{(13)(5) + (35)(47)} = \frac{65 - 1645}{65 + 1645} \\
 &= -\frac{1580}{1710} = -0.92
 \end{aligned}$$

وبالتالي يتضح أن الاقتران بين نوع الجنس وحق المرأة في العمل اقتران عكسي قوي . أي بين الذكر والرافض أو بين الأنثى والمؤيد .

ويجب ملاحظة أن معامل الاقتران يحسب في اقتران الصفة الأولى للظاهرة الأولى مع الصفة الأولى للظاهرة الثانية بالجدول ، وعليه إذا تغير وضع الجدول فإنه يغير إشارة معامل الاقتران .

9.4 معامل التوافق \bar{H} (Harmonic Coefficient)

يستخدم معامل التوافق لدراسة قوة العلاقة بين ظاهرتين وصفيتين غير خاضعتين للترتيب ولكل منهما أكثر من صفتين . وعادة ما تكون البيانات في هذه الحالة وكما هو مبين في الجدول (27-4) .

جدول (27-4)

المجموع	الظاهرة الأولى / الظاهرة الثانية		
	الصفة الأولى	الصفة الثانية	الصفة الثالثة
$\sum a_{1j}$	a_{11}	a_{12}	a_{1n}
$\sum a_{2j}$	a_{21}	a_{22}	a_{2n}
\vdots	\vdots	\vdots	\vdots
$\sum a_{nj}$	a_{n1}	a_{n2}	a_{nn}
$\sum a_{ij}$	$\sum a_{i1}$	$\sum a_{i2}$	$\sum a_{in}$

ويحسب معامل التوافق من المعادلة التالية :

$$\bar{H} = \sqrt{\frac{K' - 1}{K'}} \quad \dots\dots\dots(16-4)$$

حيث أن K' يحسب من المعادلة التالية :

$$K' = \frac{a_{11}^2}{\left(\sum_{i=1}^n a_{i1}\right)\left(\sum_{j=1}^n a_{1j}\right)} + \frac{a_{12}^2}{\left(\sum_{i=2}^n a_{i2}\right)\left(\sum_{j=2}^n a_{1j}\right)} + \dots + \frac{a_{1n}^2}{\left(\sum_{i=1}^n a_{in}\right)\left(\sum_{j=1}^n a_{1j}\right)} \quad (17-4)$$

مثال (4-14)

البيانات التالية تمثل حالة التدخين والمستوى التعليمي لثلاثمائة شخص . هل يوجد توافق بين التدخين والمستوى التعليمي لهذه البيانات والمبينة في الجدول (4-28) .

جدول (4-28)

المجموع	غير مدخن	مدخن	حالة التدخين المستوى التعليمي
			متعلم يقرأ ويكتب أمي
90	15	75	
150	60	90	
60	45	15	
300	120	180	المجموع

الحل:

نحسب معامل التوافق بالشكل التالي :

$$\begin{aligned}
 K' &= \frac{(75)^2}{(180)(90)} + \frac{(15)^2}{(120)(90)} + \frac{(90)^2}{(180)(150)} + \frac{(60)^2}{(120)(150)} \\
 &\quad + \frac{(15)^2}{(180)(60)} + \frac{(45)^2}{(120)(60)} = 1.17 \\
 \therefore \bar{H} &= \sqrt{\frac{K'-1}{K'}} = \sqrt{\frac{1.17-1}{1.17}} = \sqrt{\frac{0.17}{1.17}} = 0.38
 \end{aligned}$$

وهذا يدل على أن هناك توافقاً بين المستوى التعليمي والتدخين ولكنه ضعيف .

مثال (15-4)

الجدول (29-4) يبين لون الزهور ورائحتها فهل تعتقد أن هناك توافقاً بين لون الزهور ورائحتها ؟

جدول (29-4)

المجموع	أخضر	أبيض	أحمر	اللون ← الرائحة ↓
80	10	40	30	كوية
95	15	20	60	متوسطة
25	5	10	10	ضعيفة
200	30	70	100	المجموع

الحل :

نحسب معامل التوافق لهذه البيانات حيث نجد أن :

$$\begin{aligned}
 K' &= \frac{(30)^2}{(100)(80)} + \frac{(40)^2}{(70)(80)} + \frac{(10)^2}{(30)(80)} + \frac{(60)^2}{(100)(95)} \\
 &+ \frac{(20)^2}{(70)(95)} + \frac{(15)^2}{(30)(95)} + \frac{(10)^2}{(100)(25)} + \frac{(10)^2}{(70)(25)} \\
 &+ \frac{(5)^2}{(30)(25)} = 1.088
 \end{aligned}$$

$$\therefore \bar{H} = \sqrt{\frac{K' - 1}{K'}} = \sqrt{\frac{1.088 - 1}{1.088}} = \sqrt{\frac{0.088}{1.088}} = 0.3$$

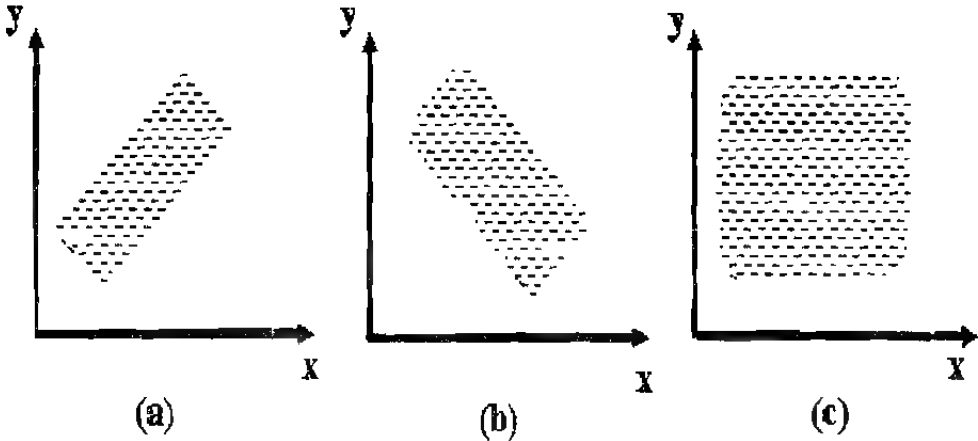
وهذا يدل على وجود توافق بين اللون والرائحة ولكنه ضعيف .

10.4 الانحدار (Regression)

لدراسة العلاقة بين ظاهرتين يمكن تكوين فكرة مبدئية عن نوع العلاقة وقوتها باستخدام ما يعرف بشكل الانتشار (Scatter Diagram) ، فإذا مثلنا أزواج المشاهدات الخاصة بالظاهرتين بيانياً نحصل على عدد من النقاط في مستوى محورين كما هو مبين في الشكل (4-1) ، حيث يتضح من الشكل (a) أنه توجد علاقة طردية بين المتغيرين ، بينما العلاقة في الشكل (b) علاقة عكسية ، ويظهر شكل الانتشار (c) أنه لا توجد علاقة بين المتغيرين حيث نجد أن النقاط مبعثرة بطريقة غير منتظمة .

وواضح من الشكلين (a) و (b) أن النقاط تقع على خط مستقيم بمعنى أنه توجد علاقة خطية بين المتغيرين يمكن وضعها في شكل معادلة من الدرجة الأولى على الصورة التالية :

$$y = m x + b \quad \dots\dots\dots (4 - 18)$$



الشكل (4 - 1)

أشكال الانتشار

حيث أن y هو المتغير التابع (Dependent Variable) ، والذي نريد تقديره و x هو المتغير المستقل (Independent Variable) ، أما m و b فهي مقادير ثابتة يمكن حسابها من واقع البيانات المشاهدة . وبمعرفة قيمة كل من m ، b يمكن استنتاج قيم y عندما نأخذ x قيمة معينة لذلك نعرف هذه المعادلة بمعادلة خط انحدار y على x حيث m تعطي ميل الخط و b هو انجزء المقطوع من المحور الرأسي .

ولتوصيل خط مستقيم يتوسط النقاط في شكل الانتشار خير توسط ليمثل العلاقة بين المتغيرين x ، y يمكننا أن نمهد هذا الخط باليد . ولكن هذا التمهيد سوف يكون تقريبياً ويختلف من شخص لآخر لذلك سنلجأ لاستخدام طريقة جبرية تعرف بطريقة المربعات الصغرى وهي طريقة دقيقة يمكننا من تحديد أفضل موضع لهذا الخط .

11.4 طريقة المربعات الصغرى (The Least Square Method)

من المعلوم أن الخط الذي نريد تمهيده سوف لا يمر بجميع النقاط في شكل الانتشار ، ولكن بعض هذه النقاط سيقع فوقه وبعضها سيقع تحته وبالتالي إذا اخترنا أي قيمة للمتغير x وقدرنا قيمة y المناظرة لها من واقع معادلة هذا الخط فإن قيمة y المقترنة سوف تختلف عن قيمة y الفعلية أي المشاهدة في حالة عدم انطباق النقطة على الخط تماماً وهذا الاختلاف يعطي لنا انحراف النقطة أي البعد الرأسي لها عن خط الانحدار .

وتهدف طريقة المربعات الصغرى إلى إيجاد معادلة لهذا الخط بحيث يكون مجموع مربعات الانحرافات أي الأبعاد الرأسية للنقط عنه أصغر ما يمكن أي ذو نهاية صغرى .

ولإيجاد معادلة هذا الخط على صورة المعادلة (4-16) حيث إن b هو الجزء المقطوع من المحور الرأسي ، و m هو ميل خط الانحدار ويسمى أيضاً بمعامل انحدار y على x نجد أن قيم m ، b والتي تحقق هذا الشرط يمكن الحصول عليها من المعادلتين :

$$\sum y = m \sum x + N b \quad \dots\dots\dots(4-19)$$

$$\sum x y = m \sum x^2 + c \sum x \quad \dots\dots\dots(4-20)$$

وبقسمة المعادلة (4-19) على N عدد المفردات نجد أن :

$$\begin{aligned} \frac{\sum y}{N} &= m \frac{\sum x}{N} + b \\ \therefore b &= \frac{\sum y}{N} - m \frac{\sum x}{N} \quad \dots\dots\dots(4-21) \end{aligned}$$

أي أن:

$$b = \bar{y} - m \bar{x} \quad \dots\dots\dots(4-22)$$

حيث أن :

- \bar{y} - هي الوسط الحسابي لقيم y .
- \bar{x} - هي الوسط الحسابي لقيم x .

وبالتعويض عن قيمة b الموجودة في المعادلة (4-19) بما يساويها في المعادلة (4-20) ينتج أن :

$$\sum x y = m \sum x^2 + \sum x \left(\frac{\sum y}{N} - m \frac{\sum x}{N} \right)$$

$$\sum x y = m \sum x^2 + \frac{(\sum x)(\sum y)}{N} - m \frac{(\sum x)^2}{N}$$

$$\therefore \sum x y - \frac{(\sum x)(\sum y)}{N} = m \left[\sum x^2 - \frac{(\sum x)^2}{N} \right]$$

وبقسمة المعادلة على N ينتج أن :

$$\frac{\sum x y}{N} - \left(\frac{\sum x}{N} \right) \left(\frac{\sum y}{N} \right) = m \left[\frac{\sum x^2}{N} - \left(\frac{\sum x}{N} \right)^2 \right]$$

$$\therefore m = \frac{\frac{\sum x y}{N} - \bar{x} \bar{y}}{S_x^2} \dots\dots\dots(23-4)$$

أي أنه يمكن معرفة كل من m ، b من المعادلتين (22-4) و (23-4) لكي نحصل على معادلة انحدار y على x .

مثال (16-4)

أوجد معادلة خط انحدار y على x من البيانات المبينة في الجدول (31-4) :

جدول (30-4)

10	2	5	4	5	6	3	x
7	1	5	6	2	4	3	y

الحل :

لإيجاد هذه المعادلة يلزم معرفة مفردات المعادلة (4-23) وهي مبينة في الجدول (4-31) :

جدول (4-31)

x y	x ²	y	x
9	9	3	3
24	36	4	6
4	25	2	5
36	16	6	4
25	25	5	5
2	4	1	2
70	100	7	10
$\sum xy = 164$	$\sum x^2 = 215$	$\sum y = 28$	$\sum x = 35$

$$\bar{x} = \frac{35}{7} = 5 \quad \bar{y} = \frac{28}{7} = 4$$

$$\therefore \sigma_x^2 = \frac{215}{7} - \left(\frac{35}{7}\right)^2 = 30.7 - 25 = 5.7$$

$$\therefore m = \frac{\frac{164}{7} - (5)(4)}{5.7} = \frac{23.43 - 20}{5.7} = \frac{3.43}{5.7} = 0.6$$

$$\therefore b = \bar{y} - m \bar{x} = 4 - (0.6)(5) = 4 - 3 = 1$$

إذن معادلة انحدار y على x هي:

$$y = 0.6x + 1$$

ويمكن كتابة المعادلة (4 - 22) الخاصة بطريقة المربعات الصغرى على الصورة التالية :

$$m = \frac{N(\sum x_n y_n) - (\sum x_n)(\sum y_n)}{N(\sum x_n^2) - (\sum x_n)^2} \dots\dots\dots(24-4)$$

$$b = \frac{(\sum x_n^2)(\sum y_n) - (\sum x_n y_n)(\sum x_n)}{N(\sum x_n^2) - (\sum x_n)^2} \dots\dots\dots(25-4)$$

حيث أن :

n تمثل رتبة القيمة في الجدول .

N = عدد القيم .

وفي هذه الحالة تستخدم المعادلة (4 - 19) لإيجاد انحدار y على x .

مثال (4-17)

تسافر سيارة على طريق مستقيم بسرعة ثابتة $V = b_1 (m/sec)$ ،

ويعطى موقع تلك السيارة y_n عند أي زمن t_n وفقاً للمعادلة $y_n = b_0 + b_1 t_n$.

أفرض أن هذه القياسات المبينة في الجدول (4-32) كانت كالتالي :

جدول (4-32)

t_n, sec	0	3	5	8	10
y_n, m	260	230	240	270	290

استخدم طريقة المربعات الصغرى لإيجاد سرعة السيارة الثابتة (b_1) .

الحل :

باستخدام المعادلات (4 - 24) و (4 - 25) ننظم الجدول (4-32) على النحو المبين في الجدول (4-33) :

جدول (4-33)

t_n	y_n	$t_n y_n$	t_n^2
0	200	0	0
3	230	690	9
5	240	1200	25
8	270	2160	64
10	290	2900	100
$\sum t_n = 26$	$\sum y_n = 1230$	$\sum t_n y_n = 6950$	$\sum t_n^2 = 198$

بما أن عدد النقاط هو 5 وباستخدام المعادلات المذكورة أعلاه نحصل على :

$$m = \frac{N(\sum x_n y_n) - (\sum x_n)(\sum y_n)}{N(\sum x_n^2) - (\sum x_n)^2} = \frac{(5)(6950) - (26)(1230)}{(5)(198) - (26)^2}$$

$$\therefore m = \frac{34750 - 31980}{990 - 676} = \frac{2770}{314} = 8.821 \frac{m}{\text{sec}}$$

$$b = \frac{(\sum x_n^2)(\sum y_n) - (\sum x_n y_n)(\sum x_n)}{N(\sum x_n^2) - (\sum x_n)^2} = \frac{(198)(1230) - (6950)(26)}{(5)(198) - (26)^2}$$

$$\therefore b = \frac{243540 - 180700}{990 - 676} = \frac{62840}{314} = 200.13$$

وبذلك تكون معادلة الانحدار هي :

$$y_m = 200.13 + 8.821 t_m$$

إن سرعة السيارة الثابتة هي :

$$8.821 \text{ m / sec}$$

مثال (4-18)

أوجد معادلة انحدار y على x من البيانات المبينة في الجدول (4-34) :

جدول (4-34)

x	5	6	7	8	9	10	11	12	13	14	15
y	45	50	53	56	60	64	70	75	78	85	91

الحل :

لاختصار العمليات الحسابية يمكن نقل نقطة الأصل وذلك بأخذ وسطين
فرضيين لقيم x و y فإذا اختيرت القيم 10 , 60 لهذا الغرض وكما هو مبين
في الجدول (4-35) ينتج أن :

$$\therefore m = \frac{\frac{493}{11} - (0)(\frac{67}{11})}{\frac{110}{11} - 0} = \frac{493}{110} = 4.482$$

$$\therefore b = \frac{67}{11} - (4.482)(0) = \frac{67}{11} = 6.091$$

$$\therefore Y = 4.482 X + 6.091$$

جدول (4-35)

x	y	X (x-10)	Y (y-60)	X ²	XY
5	45	5 -	15 -	25	75
6	50	4 -	10 -	16	40
7	53	3 -	7 -	9	21
8	56	2 -	4 -	4	8
9	60	1 -	0	1	0
10	64	0	4	0	0
11	70	1	10	1	10
12	75	2	15	4	30
13	78	3	18	9	54
14	85	4	25	16	100
15	91	5	31	25	155
		$\sum X = 0$	$\sum Y = 67$	$\sum X^2 = 110$	$\sum XY = 493$

ولحساب معادلة الانحدار بدلالة القيم الأصلية نضع $Y = y - 60$ ونضع كذلك $X = x - 10$ ونعوض في المعادلة الأخيرة أعلاه فنحصل على :

$$(y-60) = 4.482(x-10) + 6.091$$

$$\therefore y = 4.482x + 31.27$$

12.4 معادلة خط انحدار (x) على (y)

(Regression Line Equation of x on y)

إذا استخدمنا y كمتغير مستقل و x كمتغير تابع فإنه يمكن إيجاد معادلة تمكننا من تقدير قيمة x عندما تكون قيمة y معلومة وتسمى بمعادلة خط انحدار x على y ويمكن كتابتها على الصورة التالية :

$$x = \bar{m} y + d \quad \dots\dots\dots(26-4)$$

حيث أن :

$-d$ يمثل الجزء المقطوع من المحور الأفقي .

\bar{m} هي ميل خط الانحدار وتسمى أيضاً بمعامل انحدار x على y .

ويمكن إيجاد هذه باستخدام طريقة المربعات الصغرى وذلك بجعل مجموع مربعات الأبعاد الأفقية للنقاط عن خط الانحدار أصغر ما يمكن وفي هذه الحالة يمكن حساب قيم \bar{m} و d من المعادلتين التاليتين :

$$\sum x = \bar{m} \sum y + N d \quad \dots\dots\dots(27-4)$$

$$\sum yx = \bar{m} \sum y^2 + d \sum y \quad \dots\dots\dots(28-4)$$

ومن هاتين المعادلتين نجد أن :

$$\bar{m} = \frac{\sum yx - \bar{y} \bar{x}}{N \delta_y^2} \quad \dots\dots\dots(29-4)$$

$$d = \bar{x} - \bar{m} \bar{y} \quad \dots\dots\dots(30-4)$$

مثال (19-4)

أوجد معادلة خط انحدار x على y من البيانات المبينة في الجدول (36-4) .

جدول (36-4)

x	y	y^2	xy
3	3	9	9
6	4	16	24
5	2	4	10
4	6	36	24
5	5	25	25
2	1	1	2
10	7	49	70
$\sum x = 35$	$\sum y = 28$	$\sum y^2 = 140$	$\sum xy = 164$

الحل :

$$\bar{x} = \frac{35}{7} = 5 \quad \bar{y} = \frac{28}{7} = 4$$

$$\therefore S_y^2 = \frac{140}{7} - \left(\frac{28}{7}\right)^2 = 20 - (4)^2 = 20 - 16 = 4$$

$$\therefore \bar{m} = \frac{\frac{164}{7} - (5)(4)}{4} = \frac{23.43 - 20}{4} = \frac{3.43}{4} = 0.857$$

$$\therefore d = \bar{x} - \bar{m} \bar{y} = 5 - (0.857)(4) = 5 - 3.428 = 1.572$$

إن معادلة خط انحدار x على y هي:

$$x = 0.857y + 1.572$$

13.4 العلاقة بين معامل الارتباط ومعاملات الانحدار (Relation between Correlation and Regression Coefficients)

1- إن حاصل ضرب معامل انحدار y على x (m) في معامل انحدار x على y (\bar{m}) يساوي مربع معامل الارتباط أي أن :

$$m \times \bar{m} = \frac{\left(\frac{\sum xy}{N} - \bar{x} \bar{y} \right)}{S_x^2} \times \frac{\left(\frac{\sum xy}{N} - \bar{x} \bar{y} \right)}{S_y^2}$$

$$= \left(\frac{\frac{\sum xy}{N} - \bar{x} \bar{y}}{S_x^2 S_y^2} \right)^2 = R^2$$

$$\therefore R = \sqrt{m \times \bar{m}} \quad \dots\dots\dots (31-4)$$

مثال (20-4)

أحسب معامل الارتباط للبيانات المبينة في الجدول (36-4) للمثال السابق .

الحل :

$$R = \sqrt{(0.6)(0.857)} \approx 0.72$$

2- إن حاصل ضرب (m) معامل انحدار y على x في $\frac{S_x}{S_y}$ يساوي معامل الارتباط .

$$m x \frac{S_x}{S_y} = \frac{\left(\frac{\sum xy}{N} - \bar{x} \bar{y} \right)}{S_x^2} x \frac{S_x}{S_y}$$

$$= \frac{\frac{\sum xy}{N} - \bar{x} \bar{y}}{S_x S_y} = R$$

$$\therefore R = m x \frac{S_x}{S_y} \dots\dots\dots(32 - 4)$$

أو:

$$m = R x \frac{S_x}{S_y} \dots\dots\dots(33 - 4)$$

مثال (20-4)

إذا علمت أن الوسط الحسابي والانحراف المعياري للمتغير x هما 2 و 0.89 على الترتيب وأن الوسط الحسابي والانحراف المعياري للمتغير y هما 8 و 1.67 على الترتيب . أوجد معادلة خط انحدار y على x علماً أن معامل الارتباط بين قيم x ، y يساوي 0.8 .

الحل :

$$m = R x \frac{S_x}{S_y} = \frac{(0.8)(1.67)}{(0.89)} = 1.5$$

$$\therefore b = 8 - (1.5)(2) = 8 - 3 = 5$$

$$\therefore y = 1.5 x + 5$$

مثال (4-21)

إذا علمت أن معادلتى انحدار y على x وانحدار x على y هما :

$$y = 0.9x + 4.1$$

$$x = 2.1y - 1.3$$

ومن هاتين المعادلتين يتضح لنا أنه يوجد خطأ في أحدهما .

الحل :

نحسب معامل الارتباط فنجد أن :

$$R = \sqrt{m \times \bar{m}} = \sqrt{(0.9)(2.1)} = \sqrt{1.89} = 1.375$$

إنه يوجد خطأ في إحدى المعادلتين لأن معامل الارتباط لا يمكن أن يكون أكبر من الواحد الصحيح .

14.4 تمارين

س1: باستخدام طريقة المربعات الصغرى أوجد كلاً من معادلتني خط انحدار y على x ، وخط انحدار x على y للبيانات المبينة في الجدول (4-37) .

جدول (4-37)

x	1	3	4	6	8	9	11	14
y	1	3	4	4	5	7	8	9

س2: باستخدام معادلتني الانحدار في السؤال السابق أحسب معامل الارتباط بين قيم x و y .

س3: أحسب معامل ارتباط الرتب " لسبيرمان " بين قيم x و y من البيانات المبينة في الجدول (4-38) .

جدول (4-38)

x	13	14	15	11	16	12	13
y	13	16	15	12	14	15	17

س4: إذا علمت أن الانحراف المعياري لقيم x هو 0.61 والانحراف المعياري لقيم y هو 1.32 . أحسب معامل الارتباط بين قيم x و y إذا علمت أن معادلة انحدار y على x هي :

$$y = 0.84x + 21.38$$

س5: إذا علمت أن معادلتا انحدار y على x وانحدار x على y هما :

$$y = 0.72x + 3.12$$

$$x = 1.43 - 0.81y$$

بين أنه يوجد خطأ في إحدى هاتين المعادلتين .

س6: الجدول (6-39) يبين التقديرات التي حصل عليها ثمانية من الطلبة في كل من مادتي الرياضيات والإحصاء والمطلوب حساب معامل الارتباط بين تقديرات هاتين المادتين .

جدول (4-39)

تقدير الرياضيات	ضعيف	مقبول	ممتاز	جيد جداً	جيد	مقبول	مقبول	ضعيف جداً
تقدير الإحصاء	ضعيف	مقبول	جيد جداً	ممتاز	مقبول	جيد	ضعيف	ضعيف جداً

س7: أحسب معامل ارتباط بيرسون بين قيم x و y من البيانات المبينة في الجدول (4-40) .

جدول (4-40)

X	67	56	65	60	69	58
Y	177	171	170	169	174	171

أستخدم القيمتين 60 و 171 كوسطين فرضيين لقيم x و y على التوالي .

س8: أمكن التوصل إلى البيانات التالية عن المتغيرين x و y :

$$\sum x = 48 \quad , \quad \sum y = 76 \quad , \quad \sum x^2 = 536$$

$$\sum y^2 = 1108 \quad , \quad \sum xy = 720 \quad , \quad N = 6$$

والمطلوب :

(a) إيجاد معادلة انحدار y على x .

(b) إيجاد معامل الارتباط (بيرسون) بين قيم x و y .

س9: أحسب معامل الارتباط بين قيم x و y من البيانات الموضحة في

الجدول (4-41) :

جدول (4-41)

المجموع	100 - 90	- 80	- 70	- 60	- 50	X Y
2					2	- 50
20			4	11	5	- 60
35		6	25	4		- 70
30	10	14	6			- 80
13	8	5				100 - 90
100	18	25	35	15	7	المجموع

س10: أحسب معامل الارتباط العزمي ليبرسون بين قيم x , y من البيانات الموضحة في الجدول (42-4) .

جدول (42-4)

X	17 0	17 2	16 9	16 5	16 8	16 7	17 1	16 3	16 9	16 6
Y	69	71	72	66	69	66	69	67	70	69

أستخدم القيمتين 169 و 69 كوسطين فرضيين لقيم x و y على الترتيب .

س11: البيانات المبينة في الجدول (43-4) تمثل الدخل والإنفاق الشهري بالدينار لعشر أسر والمطلوب :

- 1- تحديد العلاقة المناسبة من خلال شكل الانتشار .
- 2- تقدير تلك العلاقة مع رسمها على الشكل الانتشاري .
- 3- تحديد (تقدير) الإنفاق الشهري عندما يكون الدخل 110 دينار .
- 4- تحديد معامل الارتباط بين الدخل والإنفاق .

جدول (43-4)

132	148	69	135	108	87	170	128	50	73	الدخل (x)
60	70	30	60	50	40	80	60	20	30	الإنفاق (y)

س12: البيانات المبينة في الجدول (44-4) تمثل قيمًا للمتغيرين x ، y .

جدول (44-4)

11	8	6	5	4	2	X
5	7	8	10	12	18	Y

(a) أوجد معامل ارتباط بيرسون للبيانات .

(b) أوجد معامل ارتباط الرتب لسبيرمان . أيهما أفضل مع التعليق على النتائج .

س13: من خلال البيانات أدرس العلاقة بين الإصابة بمرض الكوليرا والتطعيم بمصل مضاد من بين 200 شخص تم تطعيمهم أصيب 50 شخصاً ، ومن بين 200 شخص لم يطعموا أصيب 150 شخصاً .

س14: أحسب معامل الارتباط بين قيم x و y من البيانات المبينة (45-4) .

جدول (45-4)

المجموع	200 - 190	- 180	- 170	- 160	- 150	- 140	X Y
12				4	5	3	- 40
17			2	6	6	3	- 50
21		2	5	9	4	1	- 60
24	1	8	10	5			- 70
16	5	6	4	1			- 80
10	4	4	2				- 90
100	10	20	23	25	15	7	المجموع

س15: أدرس العلاقة بين نوع النبات ودرجة الإصابة بمرض ما من خلال البيانات المبينة في الجدول (4-46) والتي تمثل عدد النباتات المصابة من كل نوع .

جدول (4-46)

C	B	A	نوع النبات / درجة الإصابة
			بسيطة متوسطة عالية
20	30	300	
50	250	50	
300	100	30	

س16: إذا علمت أن معامل الارتباط بين درجات الكيمياء ودرجات الفيزياء مصححة من 50 هو 0.75. فإذا تم تحويل درجات المادتين من 100 فما هو معامل الارتباط الجديد .

الباب الخامس

مبادئ نظرية الاحتمالات

(Probabilities Principles)

- 1.5 مقدمة .
- 2.5 فضاء العينة .
- 3.5 النماذج الرياضية .
- 4.5 الاحتمالات القبلية .
- 5.5 الحدث .
- 6.5 احتمال الحدث .
- 7.5 التعريف التقليدي للاحتمال .
- 8.5 التعريف الإحصائي للاحتمال .
- 9.5 التعريف الحديث للاحتمال .
- 10.5 احتمال اجتماع حدثين .
- 11.5 الاحتمالات المركبة .
- 12.5 قوانين الاحتمالات .
- 13.5 الاحتمال المشروط .
- 14.5 الأحداث المستقلة .
- 15.5 التكرار النسبي والاحتمالات التجريبية .
- 16.5 التوقع الرياضي .
- 17.5 قاعدة حساب القيمة المتوقعة .
- 18.5 التحليل التوافقي .
- 19.5 التباديل والترتيبات والتوافيق .
- 20.5 تمارين .

1.5 مقدمة (Introduction)

في الأبواب السابقة تم دراسة كيفية وصف مجموعة من البيانات الإحصائية وتبويبها ، وطريق تمثيلها وعرضها بمدرجات أو مضلعات أو منحنيات تكرارية . كما تم دراسة طرق حساب المقاييس الإحصائية المختلفة وتحليلها مثل مقاييس النزعة المركزية ومنها الوسط الحسابي والمنوال والوسيط ، ومقاييس والتشتت ومنها الانحراف المعياري والتباين والانحراف المتوسط وغيرها من المقاييس .

وبعد ذلك تم التعرض لدراسة موضوع الارتباط والانحدار لمجموعة البيانات ، وبمعنى آخر ما تم دراسته وعرضه هو وصف لبيانات إحصائية متوفرة لدينا ، وهذا كله يعتبر على قدر من الأهمية لمتتبع الطرق الإحصائية . أما في هذا الباب فسوف نقوم بدراسة مبادئ نظرية الاحتمالات نظراً لارتباطها القوي والوثيق بعلم الإحصاء في الوقت الحاضر .

وتعود بدايات نظرية الاحتمالات إلى القرن السابع عشر نتيجة لدراسة بعض الألعاب الحظ المختلفة ، على الرغم أن تاريخ نظرية الاحتمالات قديم وكبير . ألا أن هذه النظرية لم توضع لها مسلمات إلا في نهاية الثلاثينات من القرن العشرين . وأصبحت تعرف على أنها العلم الذي يدرس الظواهر العشوائية ، وقد تطورت نظرية الاحتمالات تلبية لمتطلبات الحياة العملية مثلها مثل أجزاء العلوم الطبيعية المختلفة .

أن العلاقة المبنية بين نظرية الاحتمالات ومتطلبات العلوم الطبيعية الأخرى توضح بأفضل ما يمكن تلك الأسباب التي جعلت نظرية الاحتمالات في السنوات الأخيرة من أسرع فروع الرياضيات تقدماً وتقدماً . فالنتائج النظرية الجديدة تعمل على فتح آفاقاً جديدة لاستخدام طرق نظرية الاحتمالات

في العلوم الطبيعية المختلفة ، وتقوم الدراسة الشاملة للظواهر الطبيعية بدفع نظرية الاحتمالات إلى الكشف عن قوانين ومسلمات جديدة ولدت عن طريق الصدفة .

وفي السنوات الأخيرة كبرت أهمية ارتباط نظرية الاحتمالات بعلم الإحصاء بفعل التطور الصناعي والتقني السريع . ونتيجة لذلك فقد زاد الاهتمام بنتائج نظرية الاحتمالات في تنظيم عمليات الإنتاج والتصنيع ، والرقابة الإحصائية وبذات المتعلقة بمشاكل التحقق من نوعية المنتجات ، وبالتالي ظهرت نظرية الطرق الإحصائية لرقابة القبول العميقة بمحتواها ، والهامة بتطبيقاتها العملية والمبنية على الاستخدام الواسع لنظرية الاحتمالات .

ولقد كبرت أهمية نظرية الاحتمالات بكثرة في الوقت الراهن ، حيث تدخل الآن أفكار الاحتمالات والإحصاء في أغلب الاتجاهات مثل الاقتصاد والطب وعلم النفس وإدارة الأعمال والعلوم التربوية وكافة مجالات الفروع والتخصصات الهندسية . أن الهدف الرئيسي من هذا الباب هو عرض المبادئ الرئيسية لنظرية الاحتمالات ببساطة ووضوح وذلك من خلال دراسة الظواهر الطبيعية المختلفة .

فمثلاً عندما نحصل على معدل سقوط الأمطار فوق أحد الدول خلال الثلاثين سنة الماضية ، فإنه يمكننا التحدث عن تلك الثلاثين سنة الماضية ، ولكن يبقى السؤال هو ماذا يمكننا التحدث بخصوص السنوات الثلاثين القادمة . إن المعلومات التي لدينا لا تتحدث شيئاً عن السنوات القادمة ، فهل يمكننا التنبؤ بمعدل سقوط الأمطار خلال العام القادم . وأيضاً عندما نلقي قطعة من النقود مائة مرة متتالية ونجد أن عدد مرات ظهور الصورة كانت 54 مرة ، وعدد مرات ظهور الكتابة كانت 46 مرة ، فهل

يمكننا التنبؤ بشي* عن نتيجة إلقاء قطعة النقود هذه الآن . وكمثال آخر سجلت إحدى مستشفيات الولادة عدد المواليد الذكور لألف حادثة ولادة فكان العدد 495 . هل يمكن أن نقول شيئاً عن جنس المولود في حادثة ولادة جديدة ، للوصول إلى جواب حول أية من هذه الأسئلة أو أي أسئلة أخرى مماثلة يلزم دراسة مفهوم الاحتمال ونظرياته .

ويجب الإشارة إلى أنه عند اختيار عدد من العينات من " مجتمع " ما نجد أن صفات العينة تختلف من واحدة لأخرى بطريقة لا يمكن التنبؤ بها ، ونعبر عن ذلك بأن نقول إن نتائج الاختيار عشوائية أو إنها عرضة للصدفة . غير إن الباحثين قد وجدوا بالخبرة أن فكرة الاحتمالات تعطي طريقة تتضمن إجراء العديد من التجارب ذات النتائج المتوقعة على الصدفة . وقد طور علماء الرياضيات نظرية للاحتتمالات تقدم نموذجاً رياضياً (Mathematical Model) مناسباً لوصف وتقدير فئة معينة من الظواهر المشاهدة التي تقوم على أساس الصدفة .

2.5 فضاء العينة (Sample Space)

في علم الإحصاء تستخدم كلمة تجربة للدلالة على أية عملية نقودنا إلى مجموعة من البيانات . فمثلاً عند إلقاء قطعة من النقود نجد أمامنا نتيجتين محتملتين هما " ظهور الصورة " أو " ظهور الكتابة " ، وأيضاً عندما نسجل عدد الطائرات التي تهبط في مطار عمان مثلاً يومياً نجد أننا نحصل على سبعة أعداد يدل كل منها على عدد الطائرات التي هبطت في أحد أيام الأسبوع . وفي الحقيقة عندما نقوم بتجربة إحصائية ، لا نهدف إلى معرفة نتائج التجربة فحسب ، بل نتطلع من وراء هذه التجربة إلى النتائج التي نحصل عليها لو كررنا التجربة ثمانية لعدد من المرات .

إن ظهور الصورة عند إلقاء قطعة النقود لا يدل مطلقاً على أننا لو كررنا التجربة ثانية فالصورة ستظهر بالتأكيد . وإن ظهور الصورة في الحالة الأولى هو محض صدفة ، وبالتالي لا يمكن التنبؤ بظهورها ثانية بالتأكيد عند تكرار التجربة ، أو إن امرأة ما حين تضع مولوداً نكراً في الولادة الأولى لا يعني مطلقاً بأنها ستضع مولوداً نكراً في الولادة الثانية . وكل هذا يقودنا إلى تعريف فضاء العينة .

يعرف " فضاء العينة " على أنه مجموعة كل النتائج الممكنة لتجربة إحصائية ما ، ويرمز لها بالحرف S ، أي أن S هو مجموعة النتائج التي يمكن أن نحصل عليها من تنفيذ التجربة لمرة واحدة .

فمثلاً في تجربة رمي قطعة النقود لو رمزنا مثلاً بالرمز " H " لظهور الصورة وبالرمز " T " لظهور الكتابة فإن فضاء العينة لهذه التجربة يكون هو :

$$S = \{H, T\}$$

ولو رمزنا بالرمز B للمولود الذكر و الرمز G للمولود الأنثى في حالة ولادة امرأة ما ، فإن فضاء العينة لهذه التجربة يكون :

$$S = \{G, B\}$$

مثال (1-5)

نفرض أن التجربة الإحصائية هي إلقاء حجر النرد ، وأن الهدف من هذه التجربة هو الحصول على الوجه العلوي لحجر النرد ، وعليه فإن فضاء العينة يكون :

$$S = \{1, 2, 3, 4, 5, 6\}$$

أما إذا كان الهدف من التجربة الاهتمام بنوع الرقم الظاهر مثلاً زوجي (Even) أو فردي (Odd) فإن فضاء العينة عندئذ سيكون :

$$S' = \{ odd , even \}$$

إن المثال السابق يدل وجود أكثر من فضاء واحد للعينة يمكنه وصف النتائج التي يمكن عليها من خلال تكرار تجربة إحصائية معينة .

إن فضاء العينة S في المثال السابق يعطي معلومات أكثر مما قد نحصل عليه من فضاء العينة S' وهو فردي في هذه الحالة ، وذلك لأننا لو علمنا أن الرقم الذي يظهر على الوجه العلوي عند إلقاء حجر النرد هو الرقم 4 مثلاً لحكمنا على أن العنصر هو من الفئة S' أما العكس فهو غير صحيح . فلو عرفنا أن الرقم الذي ظهر على الوجه العلوي لحجر النرد كان " زوجياً " فلا يمكننا قط الجزم بأنه كان الرقم 2 أو 4 الرقم أو الرقم 6 ، لذا فإنه بصورة عامة نرغب في استخدام فضاء عينة يعطينا معلومات أوسع عن النتائج الممكنة للتجربة التي ندرسها .

مثال (2-5)

نفرض أن التجربة هي إلقاء قطعة نقود ثلاث مرات متتالية فإن فضاء العينة هو :

$$S = \{ HHH, HTH, THH, HHT, TTH, THT, HTT, TTT \}$$

ويمكن أيضاً أن يكون اهتمامنا منصّباً على عدد الصور الظاهرة خلال التجربة السابقة فيكون فضاء العينة عندئذ هو :

$$S = \{ 0, 1, 2, 3 \}$$

3.5 النمذج الرياضية (Mathematical Models)

عند استخدام الرياضيات في وصف إحدى الظواهر الطبيعية تحت الدراسة ، فإننا ننشئ نموذجاً رياضياً مبسطاً للمعالم الحقائق ويكون صورة مثالية للصفات المميزة للظاهرة التي ندرسها . ومثال على ذلك هو إدخال الإنسان فكرة النقاط والحروف والخطوط والأشكال الهندسية ومن هذه وبعض البديهيات يتكون النموذج الرياضي .

وباستخدام المنطق أستنتج علماء الرياضيات النظريات الهندسية التي ما هي إلا نظريات عن هذا النموذج الرياضي . وفي إطار هذا النموذج نجد أن النتائج الهندسية صحيحة تماماً . فمثلاً مجموع زوايا المثلث الهندسي يساوي 180° بالضبط . من المعلوم أن هذا النموذج الرياضي مفيد جداً ، ولذلك ندرس الهندسة .

وكمثال آخر نأخذ قانون نيوتن الثاني للحركة والمبني على نموذج رياضي معين . والذي ينص على أنه إذا أثرت قوة (F) على جسم كتلته (m) فإنها تزيد سرعته في اتجاه تأثيرها بمقدار (a) كل ثانية حسب العلاقة الآتية :

$$F = m a$$

ويتكون النموذج الرياضي هنا من الأعداد التي تمثل مقادير القوة (F) والكتلة (m) والعجلة (a) مع العلاقة التي تربطها . أما القانون فينص على أن هذا النموذج الرياضي يمكن أن يستعمل لوصف ظواهر طبيعية معينة . ومع أن إثبات هذا القانون رياضياً مستحيل ، ألا التجربة قد أثبتت فائدته إلى حد كبير جداً .

وبلاحظ أن قانون نيوتن الثاني للحركة يحدد قيم (a) أي معدل زيادة السرعة بالنسبة للزمن عندما تكون قيمتي القوة المؤثرة على الجسم (F) وكتلة الجسم (m) معلومة ، ولذلك نقول عن النموذج الرياضي الذي بنى قانون نيوتن على أساسه على أنه نموذج تحديدي (Deterministic) . وفي دراسة الميكانيكا والفيزياء الطبيعية (Physics) ، كثيراً ما ننشئ نماذج رياضية من هذا النوع . وعندما يكون من الضروري معرفة توزيع الحرارة المنتقلة خلال أحد الأجسام الصلبة فلا بد من الوصول إلى نموذج رياضي مناسب يوضح كيفية هذا التوزيع .

على أنه غالباً ما نقابلنا ظواهر يكون من الصعب فيها ارتباط المتغيرات مع بعضها بحيث تجعل من الصعب الحصول على نموذج رياضي تحديدي ، كالذي أستنتج في قانون نيوتن الثاني للحركة . ومع ذلك ففي مثل هذه الحالات نجد أنه من المفضل إنشاء نموذج رياضي تتوقف فيه النتيجة على الصدفة ، وبذلك نجد أنفسنا في بداية طريق يقودنا إلى ما نسميه النماذج الاحتمالية (Probability Models) .

4.5 الاحتمالات القبلية (Priority Probabilities)

عند دراسة واحداً من أبسط الأمثلة على تجارب الصدفة وهو إلقاء قطعة نقود في الهواء كما أوضحنا سابقاً . يكون من الواضح أنه لا يمكن لأحد أن يتنبأ بالنتيجة ، أن كانت هل تظهر الصورة أو الكتابة على الوجه الأعلى ، فإذا فرضنا أن القطعة سليمة وأنها ستلقى كيفما اتفق فليس هناك ما يجعلنا نتوقع ظهور أحد الوجهين أكثر مما نتوقع ظهور الوجه الآخر . وهنا ننشئ نموذجاً رياضياً تتساوى فيه فرصة ظهور الصورة والكتابة وتتصور فيه هذا التساوي للفرص بأن نخصص فيه عددين متساويين في النتيجة .

وإذا ما أخذنا بعين الاعتبار أن يكون مجموع العددين يساوي 1 فإن العدد المخصص لكل نتيجة نسميه احتمال هذه النتيجة . ويكون في هذه الحالة احتمال ظهور الصورة يساوي $\frac{1}{2}$ واحتمال ظهور الكتابة يساوي $\frac{1}{2}$ أيضاً .

وكمثال آخر على تجارب الصدفة ندرس تجربة إلقاء حجر النرد حيث يكون لدينا 6 نتائج ممكنة وهي ظهور الأعداد 1 , 2 , 3 , 4 , 5 و 6 على الوجه العلوي للنرد ، ولا يمكن بطبيعة الحال التنبؤ بالنتيجة مقدماً . ولكن إذا افترضنا أن الحجر سليم وأنه سيلقى كيفما اتفق فليس هناك ما يجعلنا نتوقع ظهور أحد الأعداد أكثر مما نتوقع ظهور العدد الآخر ، ومنه ننشئ نموذجاً رياضياً تتساوى فيه فرصة ظهور أي من الأعداد الستة 1 , 2 , 3 , 4 , 5 , 6 مع أي عدد آخر منها ونصور هذا التساوي للفرص أي نخصص أعداداً مساوية للنتائج الستة .

وإذا راعينا أن يكون مجموع هذه الأعداد يساوي 1 فإن العدد المخصص لكل نتيجة نسميه احتمال هذه النتيجة ، ويكون احتمال ظهور العدد 1 يساوي احتمال ظهور العدد 2 يساوي احتمال ظهور العدد 6 ويساوي $\frac{1}{6}$.

ولا بد أن كلاً منا قد لاحظ أن النموذج الذي أنشأناه في كل من المثالين السابقين مبني على اعتبارات لا تتضمن تجربة فعلية . ففي المثال الأول نجد أن الاحتمال يوجد قبل إلقاء قطعة النقود أي في الواقع بدون أي إلقاء لقطعة النقود . وفي المثال الثاني أوجدنا الاحتمال قبل إلقاء حجر الزهر ، والاحتمالات التي من هذا النوع تسمى الاحتمالات القبلية ومن الواضح أنها مبنية على أفكار رياضية . ولا يمكننا إلا بالتجربة وحدها إظهار فيما إذا كان للاحتمال القبلي أي معنى في مثل الحالات التي سبق الإشارة إليها .

على أننا كثيراً ما نواجه حالات لا تؤدي فيها الاعتبارات القبلية إلى تعريف كامل لنموذج احتمالي . فمن المستحيل مثلاً إيجاد احتمال أن ماكينة معينة سوف تنتج سلعة معيوبة قبل مشاهدة أداء تلك الماكينة ، وهذا ما يجعلنا ندخل نوعاً آخرًا من الاحتمالات مبنياً على اعتبارات ما بعد المشاهدة وهو ما يسمى الاحتمالات التجريبية (Empirical Probabilities) وسنقوم بدراستها لاحقاً . أما حالياً فسنقوم بدراسة الاحتمالات القبلية .

عند إجراء إحدى محاولات أو تجارب الصدفة فإن الحاصل (Result) يكون أحد النتائج الأولية (Elementary Outcomes) ، فمثلاً عندما نقوم بإلقاء حجر النرد يكون الحاصل هو 1 , 2 , 3 , 4 , 5 , 6 وعند إلقاء قطعة نقود في الهواء يكون الحاصل أحد النتيجتين الأوليتين صورة وكتابة وعند سحب ورقة من مجموعة أوراق اللعب الكوتشينة نجد أن هناك 52 نتيجة أولية يكون الحاصل واحداً منها .

مثال (5-3)

افرض أن لدينا كيساً به ثلاث كرات إحداها بيضاء والأخرى صفراء والثالثة زرقاء . وكانت التجربة تتكون من سحب كرتين عشوائياً من الكيس . أن النتائج الممكنة هي سحب الكرتين البيضاء والصفراء وسحب الكرتين الصفراء والزرقاء وسحب الكرتين الصفراء والزرقاء وإذا رمزنا للكرة البيضاء بالرمز (White) W وللكرة الصفراء بالرمز (Yellow) Y وللكرة الزرقاء بالرمز (Blue) B ، وإذا رمزنا للنتيجة الأولى على شكل زوج ثنائي مرتب هو (W,Y) وللنتيجة الثانية بالرمز (W,B) وللنتيجة الثالثة بالرمز (Y,B) فإن فضاء العينة S هو :

$$S = \{ (W,Y), (W,B), (Y,B) \}$$

5.5 الحدث (Event)

يعرف الحدث على أنه مجموعة محددة من العناصر في فضاء العينة S لتجربة من تجارب الصدفة . فمثلاً في المثال (1-5) نجد أن الحدث هو عدد زوجي يتكون من مجموعة العناصر $\{2,4,6\}$. وفي المثال (3-5) نجد أن الحدث :

$$\{(W,B),(Y,B)\}$$

هو الحدث الذي يقع عند ظهور الكرة الزرقاء في السحب .

وإذا ما أجريت التجربة وكان الحاصل إحدى النتائج الأولية التي يتكون منها الحدث لقلنا أن الحدث قد وقع أو نجح ، وألا فإننا نقول أن الحدث لم يقع أو قد فشل فمثلاً في المثال (1-5) إذا ظهر العدد 2 أو العدد 4 أو العدد 6 لقلنا أن الحدث عدد زوجي قد وقع .

ويجب أن نلاحظ أن مجموعة النتائج الممكنة التي يتكون منها الحدث يمكن أن توصف بإحدى طريقتين :

- (a) بوضعها في قائمة مثل الحدث $\{2,4,6\}$ كما في المثال (1-5) .
- (b) بنكر خاصية مشتركة بينها كأن نقول عدد زوجي في نفس المثال .

مثال (4-5)

في المثال (1-5) إذا رمزنا للعدد الذي يظهر على الوجه العلوي بالرمز x من قبل الحدث $|x > 4|$ هو الحدث $\{5,6\}$.

أن الأمثلة السابقة تقودنا إلى التعريفات الهامة التالية :

- 1- الحدث هو مجموعة جزئية من فضاء العينة .
- 2- الحدث الابتدائي هو مجموعة جزئية من فضاء العينة ذات عنصر وحيد وقد يدعى بالحدث البسيط (Simple Event) .

مثال (5-5)

في تجربة إلقاء ثلاث قطع نقود أو قطعة واحدة ثلاث مرات متتالية إذا كان الحدث A هو ظهور الصورة على إحدى القطع والكتابة على القطعتين الباقيتين حيث :

$$A = \{ HHT, THH, TTH \}$$

والحدث B هو جميع الوجوه ذات شكل واحد حيث :

$$B = \{ TTT, HHH \}$$

من المهم الملاحظة هنا أنه عند إجراء التجربة الإحصائية مرة واحدة فقط فإننا نحصل على نتيجة واحدة فقط . فمثلاً لو كانت النتيجة التي حصلنا عليها هي x فإننا نقول عن الحدث الابتدائي $\{x\}$ بأنه قد وقع ، كما نقول عن أي حادث تنتمي إليه بأنه قد وقع .

فمثلاً لو رمينا حجر النرد وظهر الرقم 3 على الوجه العلوي فإننا نقول عن الحدث :

$$A = \{ 1, 2, 3 \}$$

بأنه قد وقع .

كذلك الحال بالنسبة للحدث :

$$B = \{3, 6\}$$

لو الحدث :

$$C = \{3, 4, 5\}$$

لنفرض أن A و B حدثين موافقين لتجربة إحصائية ما . بمعنى آخر A و B فئتين جزئيتين من فضاء عينة ما S فإن A/B هو حدث أيضاً وكذلك $A \cup B$.

إن الحدث A/B هو الحدث الذي يقع إذا وفقط إذا وقع كل من A و B بينما الحدث $A \cup B$ فهو الحدث الذي يقع إذا وفقط إذا وقع أي من الحادثين A أو B أي بمعنى أبسط أن يقع أحدهما على الأقل .

مثال (5-6)

في تجربة سحب ورقة من أوراق الكوتشينة وعددها 52 ورقة . لو فرضنا أن A الحدث الذي يقع إذا كانت الورقة المسحوبة تحمل الرقم 3 . و فرضنا أن B هو الحدث الذي يقع إذا كانت الورقة المسحوبة من النوع الديناري فإن :

(A/B) - هو الحدث الذي يقع إذا كانت الورقة المسحوبة تحمل العدد 3 ونوعها ديناري .

$(A \cup B)$ - أما الحدث فهو يعني إذا كانت الورقة المسحوبة تحمل الرقم 3 أو نوعها ديناري .

ولقد أشرنا أن الحدث هو أي فئة أو مجموعة جزئية من فضاء العينة S وأن هذا الحدث يقع إذا وفقط إذا كانت نتيجة التجربة " وهو عنصر واحد من فضاء العينة " تنتمي إلى هذا الحادث وحيث إن :

$$S \subseteq S$$

فإن S هو حدث أيضاً وهو مؤكد الحدوث لأن أية نتيجة نحصل عليها من تنفيذ التجربة تنتمي إلى S وذلك حسب تعريف فضاء العينة S ، لذا فإن هذا الحدث يسمى بالحادث الأكيد أو المؤكد . كذلك نعلم بأن الفئة الخالية ϕ محتواة في أية مجموعة أي أن :

$$\phi \subseteq S$$

لذا فإنها تعتبر حدث ولكن وقوعه مستحيل لأن نتيجة التجربة لا تنتمي إلى ϕ مهما تكن هذه النتيجة ، لذلك ندعوه بالحادث المستحيل (Impossible Event) .

وهذا يقودنا إلى تعريف الإحداث المتنافية حيث أنه يقال عن الحدثان اللذان يستحيل وقوعهما معاً يدعيان "حادثين متنافيين " إذا كان :

$$A / B = \phi$$

مثال (5-7):

في تجربة رمي زهر النرد لنفرض أن الحادث A الذي يقع إذا كان الناتج رقماً زوجياً و B هو الحادث الذي يقع إذا كان الناتج عدداً فردياً فإن :

$$A = \{2, 4, 6\}$$

وأن الحدث :

$$B = \{1, 3, 5\}$$

وبالتالي فإن :

$$A/B = \phi$$

وعلى هذا الأساس فإن الحادثين A و B متنافيان .

أما الحدثان المتتامان (Complementary Events) فهما الحادثان اللذان يقع أحدهما إذا وفقط إذا لم يقع الآخر . أي أن الحدث A والحدث B متتامان إذا وفقط إذا كان :

$$A/B = \phi$$

وحيث :

$$A \cup B = S$$

أي أن :

$$B = S - A = \bar{A}$$

مثال (5-8)

الحدثان المذكوران في المثال السابق متتامان . وكذلك الحال في المثال (5-3) حيث أن الحدث M يقع إذا وفقط إذا ظهرت صورتان على الأقل والحدث N يقع إذا وفقط إذا ظهرت صورة واحدة على الأكثر هما حادثان متتامان لأن :

$$M = \{ HTH , HHT , THH , HHH \}$$

$$N = \{ HTT , THT , TTH , TTT \}$$

وواضح أن :

$$M / N = \phi$$

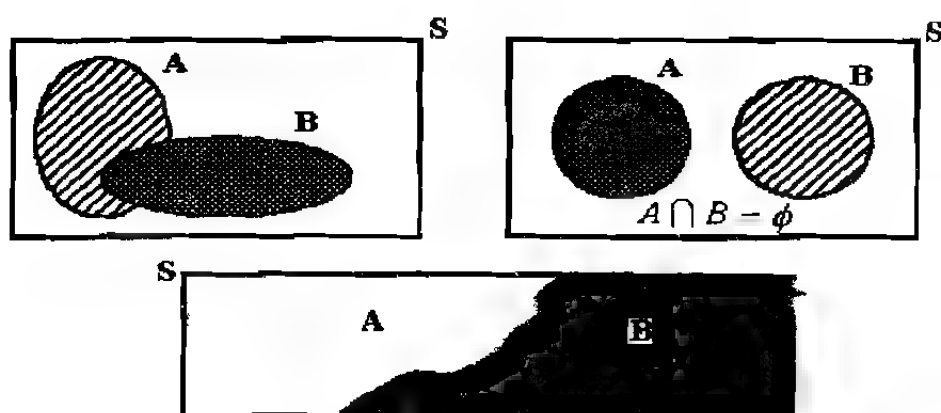
وحيث أن :

$$M \cup N = S$$

أي أن :

$$\overline{M} = N$$

ومن الممكن توضيح فضاء العينة والحوادث بأشكال فن (Venn Diagrams) ، حيث يمكن تمثيل فضاء العينة S بمستطيل وتمثيل الحوادث بنوائر أو أشكال هندسية أخرى منتظمة مثلث ، شبه منحرف وغيرها ، واقعة داخل هذا المستطيل أو بمناطق جزئية من هذا المستطيل كما يوضح الشكل (1-5) .



الشكل (1-5)

تمثيل الحوادث بأشكال فن

6.5 احتمال الحدث (Probability of the Event)

لقد أشرنا سابقاً أن الباحث الإحصائي يهتم عادةً بالوصول إلى قرارات أو استنتاجات من التجارب التي يجريها ، ويهتم بأن يكون استنتاجه أو قراره معقولاً إلى درجة ما ، لذا ينبغي أن يلم بالقواعد الأساسية لمفهوم الاحتمال .

ولأجل ذلك ندرس العبارات التي عادة ما نسمعها يومياً مثل احتمال فوز فريق بكرة القدم على فريق آخر ضمن بطولة الدوري هو 0.7 . واحتمال أن يكون المولود القادم ذكراً هو 0.6 ، أو احتمال أن يتعين في وظائف أكثر من نصف خريجي العام الحالي هو 0.4 .

إن كل من العبارات السابقة تعبر عن نتيجة غير مؤكدة لتجربة إحصائية معينة ، لكن فهمنا لطبيعة التجربة واعتمادنا على معلومات سابقة مكونة من دراستنا لمجموعة بيانات إحصائية حول مثل هذه التجربة يجعلنا على درجة من اليقين تتيح لنا الحكم بالنتيجة المقبلة . لذا سنقوم بذكر ثلاث تعاريف لمفهوم الاحتمال نعتمد جميعها على مفهوم فضاء العينة S والحدث A .

7.5 التعريف التقليدي للاحتمال (Probability's Classical Definition)

أشرنا سابقاً أن نظرية الاحتمالات هي نظرية دراسة التجارب العشوائية وقد بدأت من الناحية التاريخية بدراسة بعض ألعاب الحظ مثل الروليت والورق ، فإذا رمي حجر النرد في الهواء فمن المؤكد أنه سوف يسقط على الأرض ولكن ليس من المؤكد مثلاً أن العدد 6 سوف يظهر ، ولكن إذا فرضنا أننا كررنا هذه التجربة في رمي حجر النرد وأن S هو عدد مرات النجاح أي

عدد مرات ظهور العدد 6 ، وأن n هو عدد رميات حجر النرد فقد لوحظ تجريبياً أن النسبة $f = S / n$ والتي تسمى بالتكرار النسبي تصبح مستقرة في المدى الطويل أي أنها تقترب من نهاية ما ، حيث يعتبر هذا الاستقرار هو أساس نظرية الاحتمالات .

ولقد عرف الاحتمال P لحدث A كلاسيكياً كما يلي : أن كان الحدث A يمكن أن يقع بطرق عددها n من بين طرق كلية عددها s بشرط أن تكون لهذه الطرق نفس الفرصة في الوقوع فإن :

$$P(A) = \frac{s}{n}$$

فمثلاً عند ألغاء حجر النرد فإن الأعداد الفردية يمكن أن تقع بثلاث طرق من ست طرق لها نفس فرصة الوقوع ، إذ أن $P = 3/6 = 1/2$. وهذا التعريف التقليدي أو الكلاسيكي للاحتمال هو بالضرورة تعريف دائري ، إذ أن تعبير له نفس فرصة الظهور هو تعبير له نفس الاحتمال .

وعليه عندما تكون النتائج الأولية لإحدى المحاولات أو تجارب الصدفة محدودة العدد ونكون متفقين على أنها متساوية الوقوع (Equally Likely) أي يكون لها نفس الفرصة في الوقوع فإننا نخصص لكل حدث مثل (A) عدداً نسبيته احتمال وقوع أو نجاح A كما يلي :

$$\text{احتمال وقوع } (A) = \frac{\text{عدد النتائج الأولية التي يتكون منها}}{\text{عدد النتائج الأولية في التجربة}}$$

لو :

$$P(A) = \frac{M}{M+N} \dots\dots\dots(1-5)$$

حيث إن :

$P(A)$ - احتمال وقوع الحدث (A)

M - عدد النتائج الأولية التي يتكون منها الحدث (A) .

N - عدد النتائج الأولية التي لا يتكون منها الحدث (A) .

أما الحدث الذي يتكون من النتائج الأولية التي لا تدخل في (A) ، فإننا نسميه الحدث (\bar{A}) ومن الواضح أن وقوعه يعني الفشل أو عدم حصول الحدث وبذلك يكون :

$$\text{احتمال وقوع } (\bar{A}) = \frac{\text{عدد النتائج الأولية التي لا تدخل في } A}{\text{عدد النتائج الأولية في التجربة}}$$

لو :

$$P(\bar{A}) = \frac{N}{M+N} \dots\dots\dots(2-5)$$

ومن تعريف احتمال وقوع (A) واحتمال عدم وقوعه نجد المسلمة التالية :

$$P(A) + P(\bar{A}) = 1 \dots\dots\dots(3-5)$$

ومن هذه المسلمة نحصل على النتيجة التاليتين :

النتيجة الأولى : احتمال وقوع حادث مستحيل الوقوع يساوي صفر .

النتيجة الثانية : احتمال وقوع حادث مؤكد الوقوع يساوي الواحد .

مثال (5-9)

عند سحب ورقة لعب من أوراق الكوتشينة التي عددها 52 فإن احتمال وقوع الحادث A الذي يقع إذا فقط إذا كانت الورقة المسحوبة دينارية هو :

$$P(A) = \frac{13}{52} = \frac{1}{4}$$

كما أن احتمال وقوع الحادث B الذي يقع إذا فقط إذا كانت الورقة المسحوبة تحمل الرقم 5 هو:

$$P(B) = \frac{4}{52} = \frac{1}{13}$$

أما احتمال أن لا تكون الورقة المسحوبة تحمل الرقم 5 :

$$P(\bar{B}) = \frac{40}{52} = \frac{12}{13}$$

8.5 التعريف الإحصائي للاحتمال

(Statistical Definition of Probability)

أن ما يعيب التعريف الكلاسيكي التقليدي للاحتمال أن عبارة " كل منها له نفس الفرصة في الظهور " هي عبارة غامضة نوعاً ما ، وقد تبدو أنها مرادفة للعبارة " كل منها له نفس الاحتمال " وبذلك نكون قد عرفنا الاحتمال بدلالة نفسه . ولهذا فإن البعض يلجأ عادةً إلى إعطاء الاحتمال تعريفاً إحصائياً بحثاً .

وهذا كله ما يقودنا إلى التعريف الذي أشرنا إليه سابقاً وهو إن احتمال حدث ما هو التكرار النسبي لوقوع هذا الحدث عندما يكون عدد المشاهدات كبيراً جداً أي أنه نهاية التكرار النسبي لوقوع هذا الحادث عندما يؤول عدد المشاهدات إلى المالانهاية .

مثال (5-10)

إذا ألقيت قطعة نقود 1000 مرة وحصلنا على 529 صورة فإن التكرار النسبي للصورة هو :

$$\frac{529}{1000} = 0.529$$

ثم إذا قذفنا قطعة النقود 1000 مرة أخرى وحصلنا على 493 صورة فإن التكرار النسبي للصورة في 2000 رمية هو :

$$\frac{493 + 529}{2000} = 0.511$$

وإذا قمنا بتكرار التجربة مرات ومرات فإننا نقرب أكثر فأكثر من عدد نسميه احتمال ظهور الصورة في رمية واحدة لقطعة النقود .

ونلاحظ هنا أننا نقرب من العدد 0.5 ، بعد إجراء كل تجربة ، ولكن للوصول إلى عدد دقيق يجب إعادة التجربة مرات ومرات .

9.5 التعريف الحديث للاحتمال

(The Modern Definition of Probability)

إن التعريف الإحصائي يبدو مفيداً من الناحية العملية ، لكنه صعب التحقيق من وجهة النظر الرياضية لأنه يتطلب تكرار التجربة عدداً كبيراً جداً من المرات للوصول إلى نهاية للتكرار النسبي وقد لا نصل أحياناً إلى هذه النهاية . كما إن اعتماد التعريف التقليدي يتطلب فضاء عينة عناصرها كلها متساوية الاحتمال .

إن التعريف الحديث للاحتمال هو فرضي بحت (Purely Axiomatic) حيث يجب أن نذكر أولاً تعريف التابع الاحتمالي (Probability Function) لفضاء عينة فئة . فمثلاً لتكن S فضاء عينة ما ولتكن P هي دالة من مجموعة أجزاء من S إلى مجموعة الأعداد الحقيقية R فإن P يسمى تابعاً احتمالياً إذا حققت الشروط التالية :

مهما يكن :

$$A \leq S$$

فإن :

$$0 \leq P(A) \leq 1$$

وإن :

$$P(S) = 1$$

من أجل حدثين متناقضين مثل A و B فإن :

$$P(A \cup B) = P(A) + P(B) \dots \dots \dots (4 - 5)$$

وذلك بفرض S فضاءاً منتهياً ، أي ذو عدد محدد من العناصر ، واعتماداً على التعريف السابق وبفرض S ذات عدد n من العناصر وهي x_1, x_2, k, x_n ، ويمكن أن نستنتج أن :

$$P(S) = \sum_{i=1}^n P(x_i) \dots\dots\dots (5-5)$$

فإذا كانت جميع الحوادث الابتدائية $\{x_i\}$ متساوية الاحتمال فإن :

$$P(x_i) = \frac{1}{n} \dots\dots\dots (6-5)$$

وبالتالي فإن احتمال حدث A عدد عناصره k هو :

$$P(A) = \frac{k}{n}$$

كما ينتج مباشرة من هذا التعريف أن :

$$P(\phi) = 0$$

مثال (5-11)

عند إلقاء قطعة نقود متوازنة مرتين متتاليتين أوجد احتمال ظهور صورة واحدة على الأقل .

الحل :

بما أن قطعة النقود متوازنة فإن احتمال ظهور الكتابة وظهور الصورة متساويا الإمكانية في الظهور أي إن فضاء العينة هو :

$$S = \{ HH, HT, TH, TT \}$$

وعناصر S كلها متساوية الاحتمال ، إن الحدث الذي يقع إذا ظهرت صورة واحدة على الأقل هو :

$$A = \{ HH, HT, TH \}$$

وبالتالي فإن :

$$P(A) = \frac{3}{4}$$

مثال (12-5)

عائلة لديها ثلاثة أطفال ، ما احتمال أن يكون عددهم صبيان وبنت علماً أن للبنات وللصبي نفس الفرصة بالولادة .

الحل :

نقوم بوضع فضاء العينة S حيث :

$$S = \{ GGG, GGB, GBG, BGG, GBB, BGB, BBG, BBB \}$$

إن الحدث المطلوب هو :

$$A = \{ GBB, BBG, BGB \}$$

وبالتالي احتمال أن يكون لهذه العائلة صبيان وبنت :

$$P(A) = \frac{3}{8}$$

مثال (5-13)

ألقي زهر نرد بطريقة ما بحيث أن فرصة ظهور العدد الزوجي تساوي ضعف فرصة ظهور العدد الفردي . ما هو احتمال الحدث A الذي يقع إذا وفقط إذا كان الناتج أقل من 4 ، وما هو احتمال الحدث B الذي يقع إذا وفقط إذا كان الناتج يقبل القسمة على 3 ، ثم ما هو احتمال الحدث C الذي يقع إذا وفقط إذا كان الناتج عدداً زوجياً .

الحل :

إن فضاء العينة S هو :

$$S = \{ 1, 2, 3, 4, 5, 6 \}$$

فإذا فرضنا أن فرصة ظهور العدد الفردي هي x ، فإن احتمال ظهور العدد الزوجي ستكون $2x$ وبالتالي فإن :

$$P(S) = P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) + P(\{5\}) + P(\{6\})$$

$$1 = x + 2x + x + 2x + x + 2x \Rightarrow \therefore 1 = 9x \rightarrow \therefore x = \frac{1}{9}$$

وعلى هذا الأساس فإن الحدث هو :

$$A = \{ 1, 2, 3 \}$$

وااحتماله هو :

$$P(A) = P(\{1\}) + P(\{2\}) + P(\{3\})$$

$$= \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$$

أما الحدث B أي ظهور عدد يقبل القسمة على ثلاثة فهو :

$$B = \{ 3, 6 \}$$

واحتماله هو:

$$\begin{aligned} P(B) &= P(\{3\}) + P(\{6\}) \\ &= \frac{1}{9} + \frac{2}{9} = \frac{3}{9} = \frac{1}{3} \end{aligned}$$

والحدث C أي إذا كان الناتج عدد زوجي هو :

$$C = \{ 2, 4, 6 \}$$

واحتماله هو :

$$\begin{aligned} P(C) &= P(\{2\}) + P(\{4\}) + P(\{6\}) \\ &= \frac{2}{9} + \frac{2}{9} + \frac{2}{9} = \frac{6}{9} = \frac{2}{3} \end{aligned}$$

10.5 احتمال اجتماع حدثين

(Probability of the Union of Two Events)

مبرهنة 1:

ليكن A و B حدثين غير متنافيين أي أن $A \cap B \neq \emptyset$ فإن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \dots\dots\dots(7-5)$$

البرهان :

إن الحدث A والحدث $B - (A/B)$ حدثان متنافيان فإن :

$$A \cup B = A \cup [B - (A/B)]$$

وبالتالي فإن:

$$(1) P(A \cup B) = P(A) + P[B - (A/B)]$$

ولكن A/B و $B - (A/B)$ حدثان متنافيان :

$$B = (A/B) \cup [B - (A/B)]$$

إذن :

$$(2) P(B) = P(A/B) + P[B - (A/B)]$$

$$\therefore P[B - (A/B)] = P(B) - P(A/B)$$

من (1) و (2) نستنتج أن :

$$P(A \cup B) = P(A) + P(B) - P(A/B)$$

مثال (5 - 14)

سحبت ورقة لعب من شدة عدد أوراقها 52 . ما هو احتمال أن تكون
دينارية أو ملك .

الحل :

ليكن A الحدث الذي يقع إذا فقط إذا كانت الورقة المسحوبة دينارية ، و B الحدث الذي يقع إذا فقط إذا كانت الورقة المسحوبة هي ملك وبالتالي فإن A/B هو الحدث الذي يقع إذا فقط إذا كانت الورقة المسحوبة ملك ديناري .

أن احتمال الحدث A هو :

$$P(A) = \frac{13}{52} = \frac{1}{4}$$

وا احتمال الحدث B هو :

$$P(B) = \frac{4}{52} = \frac{1}{13}$$

وأن :

$$P(A/B) = \frac{1}{52}$$

وعلى هذا الأساس فإن :

$$P(A \cup B) = P(A) + P(B) - P(A/B)$$

$$\therefore P(A \cup B) = \frac{1}{4} + \frac{1}{13} - \frac{1}{52} = \frac{13+4-1}{52} = \frac{16}{52} = \frac{4}{13}$$

11.5 الاحتمالات المركبة (Compound Probabilities)

في كثير من الأحيان قد تتكون التجربة من تجربتين بسيطتين ، فمثلاً قد نقوم عند إجراء التجربة البسيطة رمي قطعة نقود ، بأجراء التجربة البسيطة رمي حجر النرد ، وبذلك نحصل على التجربة المركبة أو الاحتمالية المركبة وهي " رمي قطعة نقود ورمي حجر النرد بنفس الوقت " ، وإذا رمزنا للتجربة البسيطة الأولى بالرمز C_1 ، والتجربة البسيطة الثانية بالرمز C_2 فإن التجربة المركبة تكون هي الزوج المرتب (C_1, C_2) . وهنا يجب اتباع القواعد التالية :

القاعدة الأولى :

إذا أمكن أن يكون للتجربة C_1 نتائج عددها n_1 ، والتجربة C_2 نتائج عددها n_2 ، فإنه يكون للتجربة المركبة C_1, C_2 نتائج عددها $n_2 \times n_1$. فمثلاً للتجربة " رمي قطعة نقود " نتيجتان هما " صورة وكتابة " وللتجربة رمي زهر النرد 6 نتائج فإذا رمزنا للصورة بالرمز x والكتابة بالرمز y فإن التجربة المركبة نتائج عددها 6×2 وهي :

$(x,1), (x,2), (x,3), (x,4), (x,5), (x,6)$

$(y,1), (y,2), (y,3), (y,4), (y,5), (y,6)$

حيث أن $(x,1)$ تعني الصورة على ظهر قطعة النقود والعدد 1 على الحجر ، وهكذا $(y,2)$ يعني ظهور الكتابة على ظهر قطعة النقود والعدد 2 على الحجر . كذلك إذا ألقى حجراً نرد زهر معاً فإن لهذه التجربة المركبة نتائج عددها 36 وهي موضحة في الجدول (1-5) .

جدول (1-5)

نتائج التجربة المركبة من رمي حجرى نرد زهر

(1 , 6)	(1 , 5)	(1 , 4)	(1 , 3)	(1 , 2)	(1 , 1)
(2 , 6)	(2 , 5)	(2 , 4)	(2 , 3)	(2 , 2)	(2 , 1)
(3 , 6)	(3 , 5)	(3 , 4)	(3 , 3)	(3 , 2)	(3 , 1)
(4 , 6)	(4 , 5)	(4 , 4)	(4 , 3)	(4 , 2)	(4 , 1)
(5 , 6)	(5 , 5)	(5 , 4)	(5 , 3)	(5 , 2)	(5 , 1)
(6 , 6)	(6 , 5)	(6 , 4)	(6 , 3)	(6 , 2)	(6 , 1)

حيث أن x , y تعني ظهور العدد x على الزهر الأول والعدد y على الزهر الثانى .

القاعدة الثانية :

إذا كان للتجربة C_1 نتائج أولية متساوية فرصة الوقوع عددها n_1 وللتجربة C_2 نتائج متساوية فرصة الوقوع عددها n_2 فإن النتائج الأولية للتجربة المركبة " C_2 , C_1 " والتي عددها $n_2 \times n_1$ تكون أيضا متساوية في فرص الحدوث والوقوع .

مثال (5 - 15)

يلقى لاعب حجرى نرد معاً مرة واحدة . أوجد احتمال أن يحصل على مجموع أقل من 4 أو أكثر من 10 .

الحل :

يكون المجموع أقل من 4 أو أكثر من 10 إذا ظهر على أحد الوجهين العلويين للحجرين عدنان كما في أحد الحالات الستة المبينة في الجدول (2-5) .

الجدول (2-5)

6	6	5	2	1	1	العدد على الحجر الأول
6	5	6	1	2	1	العدد على الحجر الثاني
12	11	11	3	3	2	المجموع

عدد النتائج الأولية للتجربة يساوي $6 \times 6 = 36$ ، ومن هذه النتائج المتساوية الإمكان حسب القاعدة السابقة توجد 6 احتمالات ينجح فيها الحادث " المجموع أقل من 4 أو أكثر من 10 " . ومنه يكون الاحتمال المطلوب يساوي :

$$\frac{6}{36} = \frac{1}{6}$$

12.5 قوانين الاحتمالات (Probabilities Laws)

1.12.5 قانون جمع الاحتمالات (Addition Law)

تعريف

يقال للحدين A و B أنهما مانعان بالتبادل (Mutually Exclusive) إذا كان وقوع أحدهما يتنافى مع وقوع الآخر . ومن الواضح أن هذا يعني أن

الحدث A والحدث B يسميان مانعين بالتبادل إذا كانت النتائج التي يتكون منها A لا تتداخل في النتائج التي يتكون منها B .

فمثلاً إذا ألقى زهر النرد فإن النتائج (3,1) لا تتداخل مع النتائج {6,4,2} لذا فإن الحدث {odd number} أو عدد فردي والحدث {even number} أو عدد زوجي مانعان بالتبادل .

ومن الواضح أنه إذا كان A و B مانعان بالتبادل فإن عدد النتائج في الحدث $A \cup B$ أي $A \vee B$ ، تعني A أو B وهي أحد أدوات الربط المنطقية يساوي مجموع عدد النتائج في A وعدد النتائج في B . ففي مثال إلقاء حجر النرد نجد أن عدد نتائج الحادث فردي أو زوجي يساوي $3 + 3 = 6$ ومن هذا سوف ينتج أن :

$$P(A \cup B) = P(A) + P(B) \quad \dots\dots\dots (5-8)$$

أما إذا كان الحدثين غير مانعين بالتبادل فإن :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \dots\dots\dots (5-9)$$

13.5 الاحتمال المشروط (Conditional Probability)

ذكرنا أنه توجد في أساس احتمال مجموعة معينة من الظروف . وأن لم تكن هناك أية ظروف أخرى عند حساب الاحتمال لحدث ما مثل A ما عدا الظروف الأساسية المذكورة سابقاً فإن هذه الاحتمالات تسمى عندئذ بالاحتمالات اللاشرطية (Unconditional Probability) .

ولكن نضطر في بعض الحالات إلى إيجاد الاحتمال بشرط إضافي ينحصر في وقوع الحدث B . وتسمى مثل هذه الاحتمالات بالاحتمالات الشرطية (Conditional probability) ، ويرمز لها بالرمز $P(A / B)$ وهذا يعني احتمال الحدث A بشرط وقوع الحدث B . وفي الحقيقة عند الحديث بصورة دقيقة ، فإن الاحتمالات اللاشرطية هي أيضاً احتمالات شرطية لأنه عند بناء نظرية الاحتمالات أفترض وجود مجموعة معينة ثابتة من الشروط .

لنفرض أن شخصاً ألقي زهر نرد ثم أخبرنا أن العدد الذي ظهر زوجي ، فما هو احتمال أن يقبل هذا العدد القسمة على 3 . لنفرض أن الحدث A هو الحدث الذي يمثل {عدد يقبل القسمة على 3} ، وليكن B هو الحدث الذي يمثل {عدد زوجي} ، والسؤال المطروح ما هو احتمال وقوع الحادث A مع العلم بأن B قد وقع .

وسنرمز لهذا الاحتمال بالرمز $P(A/B)$ ويعني احتمال A بفرض أن الحدث قد وقع B ، أو بعبارة أخرى الاحتمال المشروط للحدث A عند وقوع B ونسميه احتمالاً شرطياً .

فمثلاً عند إلقاء حجر النرد نجد أن الحدث {عدد زوجي} له النتائج 2, 4, 6 وعددها 3 ، والحدث {يقبل القسمة على 3} له النتائج 3, 6 .

إذن الحادث {عدد زوجي ويقبل القسمة على 3} له النتيجة 6 وعددها 1 ، فعلى فرض تساوي فرص الوقوع فإن :

$$P(\text{زوجي} | \text{يقبل القسمة على 3}) = \frac{\text{عدد نتائج } \{\text{عدد زوجي ويقبل القسمة على 3}\}}{\text{عدد نتائج } \{\text{عدد زوجي}\}} = \frac{1}{3}$$

وبصورة عامة فإنه إذا كان A و B هما حادثين غير مانعين بالتبادل أو غير متنافيين كما تم الإشارة سابقاً فإن :

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \dots\dots\dots(10-5)$$

$$\therefore P(A \cap B) = P(B).P(A/B) \dots\dots\dots(11-5)$$

مثال (5-16)

توجد ثلاثة أكياس متشابهة بها كرات بعضها أبيض والآخر أسود . في الكيس الأول 30 كرة وفي الثاني 20 كرة وفي الثالث 10 كرات . عدد الكرات البيضاء في كل كيس 5 بالضبط . وقد اختير أحد الأكياس كيفما اتفق وسحبت منه كرة . أوجد احتمال أن تكون الكرة المسحوبة من الكيس الأول وتكون بيضاء في نفس الوقت .

الحل:

احتمال أن تكون الكرة من الكيس الأول يساوي $\frac{1}{3}$ ، وذلك بفرض تساوي فرص السحب لكل من الأكياس الثلاثة . أما احتمال سحب كرة بيضاء من الكيس الأول فيساوي :

$$\frac{5}{30} = \frac{1}{6}$$

على هذا الأساس فإن احتمال أن تكون الكرة المسحوبة من الكيس الأول وبيضاء هو :

$$\frac{1}{6} \times \frac{1}{3} = \frac{1}{18}$$

مثال (5-17)

في إحدى الكليات الهندسية وجد أن 25% من الطلبة رسبوا في مادة مقاومة المواد ، ووجد أيضاً أن 15% من الطلبة رسبوا في مادة الرياضيات ، وأن 10% رسب في مائتي مقاومة المواد والرياضيات . إذا اختير أحد الطلبة بطريقة عشوائية فأوجد ما يلي :

- 1- احتمال أن يكون راسباً في الرياضيات إذا كان راسباً في مقاومة المواد .
- 2- احتمال أن يكون راسباً في مقاومة المواد إذا كان راسباً في الرياضيات .
- 3- احتمال أن يكون راسباً في المائتين أي في مقاومة المواد والرياضيات .

الحل :

نفرض أن الحدث A هو الرسوب في مادة مقاومة المواد . والحدث B هو الرسوب في مادة الرياضيات ، فيكون لدينا :

$$P(A)=0.25$$

$$P(B)=0.15$$

$$P(A \cap B)=0.10$$

أولاً : احتمال أن يكون الطالب راسباً في مقاومة المواد علماً بأنه راسباً في الرياضيات هو احتمال مشروط يمكن الحصول عليه كما يلي :

$$\begin{aligned} P(A/B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{0.10}{0.15} = \frac{2}{3} \end{aligned}$$

ثانياً : احتمال أن يكون الطالب راسباً في الرياضيات علماً بأنه راسب في مادة مقاومة المواد هو :

$$\begin{aligned} P(B / A) &= \frac{P(B \cap A)}{P(A)} \\ &= \frac{0.10}{0.25} = \frac{2}{5} \end{aligned}$$

ثالثاً : أما احتمال أن يكون راسباً في المادتين فهو :

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.25 + 0.15 - 0.10 \\ &= 0.3 \end{aligned}$$

أن هذا المثال يقودنا إلى تعريف نظرية الضرب للاحتمال المشروط ، ففي العلاقة المستخدمة في حل المثال السابق نجد أن :

$$P(A / B) = \frac{P(A \cap B)}{P(B)}$$

إذا ضربنا الطرفين في الوسطين نحصل على العلاقة التالية :

$$P(A \cap B) = P(B) P(A / B)$$

وتسمى هذه العلاقة بنظرية الضرب للاحتمال المشروط ، ويمكن تعميم هذه النظرية بالاستنتاج الرياضي التالي :

لأي مجموعة من الإحداث مثل A_1, A_2, \dots, A_n :

$$P(A_1 \cap A_2 \cap \dots \cap A_n) \\ = P(A_1)P(A_2 / A_1)P(A_3 / A_1 \cap A_2) \dots P(A_n / A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

مثال (5-18)

صندوق يحتوي على 12 كرة سلة ، منها 4 كرات معيبة ، إذا اختيرت بطريقة عشوائية ثلاث كرات من هذا الصندوق واحدة تلو الأخرى ، أوجد الاحتمال P أن تكون الثلاث كرات سليمة .

الحل :

أن احتمال أن تكون الكرة الأولى سليمة يساوي $8/12$ ، حيث أن هناك 8 كرات سليمة من بين 12 كرة موجودة في الصندوق . إذا كانت الكرة الأولى سليمة فإن احتمال أن تكون الكرة الثانية سليمة يساوي $7/11$ حيث أن هناك 7 كرات فقط سليمة من بين 11 كرة باقية في الصندوق .

إذا كانت الكرة الأولى والكرة الثانية سليمتين فإن احتمال أن تكون الكرة الأخيرة سليمة هو $6/10$ حيث أن هناك 6 كرات فقط سليمة من بين 10 وحدات باقية في الصندوق . أنن باستخدام نظرية حاصل الضرب للاحتمال المشروط نجد أن :

$$P = \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} = \frac{14}{55}$$

أن دراسة الاحتمال المشروط ونظرية الضرب للاحتمال المشروط تقودنا إلى دراسة التجزيئات ونظرية "العالم ببيتز" والتي تتضمن ما يلي :

نفرض أن الأحداث A_1, A_2, \dots, A_n تمثل مجموعات جزئية لفضاء العينة S ، أي أن الأحداث متنافية واتحادهما يعطينا فضاء العينة S ، ونفرض أن الحدث B هو أي حدث آخر فيكون :

$$\begin{aligned} B &= S \cap B = (A_1 \cup A_2 \cup \dots \cup A_n) \cap B \\ &= (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_n \cap B) \end{aligned}$$

وبما أن الأحداث $A_i \cap B$ متنافية أنن :

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \dots + P(A_n \cap B)$$

وباستخدام نظرية حاصل الضرب للاحتمال المشروط نجد أن :

$$P(B) = P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + \dots + P(A_n)P(B/A_n)$$

ومن ناحية أخرى لكل i يكون الاحتمال المشروط للحدث A_i عند وقوع B هو :

$$P(A_i/B) = \frac{P(A_i \cap B)}{P(B)}$$

وللوصول إلى نظرية العالم ببيتز سوف نقوم بالتعويض في هذه المعادلة عن $P(B)$ من العلاقة السابقة ونعوض عن $P(A_i \cap B)$ بالمقدار $P(A_i)P(B/A_i)$.

نظرية بيز :

نفرض أن الأحداث A_1, A_2, \dots, A_n تمثل مجموعات جزئية لفضاء العينة S ، أي تجزئي لفضاء العينة S ، ونفرض أن الحدث B هو أي حدث آخر فيكون لكل i :

$$P(A_i / B) = \frac{P(A_i) P(B / A_i)}{P(A_1) P(B / A_1) + P(A_2) P(B / A_2) + \dots + P(A_n) P(B / A_n)}$$

أن هذه النظرية تعتبر من النظريات والقواعد الهامة في علم الاحتمالات . والمخطط العام لاستخدام هذه النظرية في حل المسائل العملية هو أن نفرض أنه يمكن أن يجري الحدث B في ظروف مختلفة ، ويمكن لطبيعتها وضع n من الفرضيات A_1, A_2, \dots, A_n . ولسبب من الأسباب عرفنا الاحتمالات $P(A_i)$ ، ونعلم أيضا أن الفرضية A_i تكسب الحدث B احتمالا مقداره $P(B / A_i)$. وأجريت تجربة وقع فيها الحدث B فإن هذا يقودنا في النهاية إلى إعادة تقدير احتمالات الفرضيات A_i ، حيث تقوم نظرية بيز بحل هذه المسألة من الناحية الكمية وأن الأمثلة الآتية سوف توضح لنا أهمية هذه النظرية في التطبيقات العملية .

مثال (5-19)

في أحد المراكز الخاصة بالتدريب البدني وجد أن 4% من الطلبة الرجال و 1% من الطلبة النساء أطول من 1.75m . وكان 60% من طلبة هذا المركز من النساء . فإذا اختير أحد الطلبة بطريقة عشوائية ووجد أنه أطول من 1.75m . أوجد احتمال أن يكون ها الطالب امرأة .

الحل :

نفرض أن الحدث A هو أطول من 1.75m ، والمطلوب هو إيجاد احتمال أن يكون الطالب المختار امرأة بمعلومية أنه أطول من 1.75m أي المطلوب إيجاد $P(W/A)$. من نظرية بيزر يمكننا إيجاد الاحتمال المطلوب حيث :

$$\begin{aligned} P(W / A) &= \frac{P(W)P(A / W)}{P(W)P(A / W) + P(M)P(A / M)} \\ &= \frac{(0.60)(0.02)}{(0.60)(0.02) + (0.40)(0.40)} \\ &= \frac{3}{11} = 0.27 \end{aligned}$$

وهذا هو احتمال أن يكون الطالب المختار بطريقة عشوائية امرأة وأطول من 1.75m .

مثال (20-5)

في شركة صناعية كبرى متخصصة بصناعة الرقائق الكترونية تنتج ثلاث الآلات A , B , C على التوالي 60% ، 30% ، 4% من الإنتاج الكلي للشركة . فإذا كان نسبة إنتاج الرقائق المعيبة لهذه الآلات هي على التوالي 2% ، 3% ، 4% . فإذا اختيرت رقيقة الكترونية بطريقة عشوائية ووجدت أنها معيبة . أوجد احتمال أن تكون هذه الرقيقة من إنتاج الآلة C .

الحل :

نفرض أن الحدث M هو أن تكون الرقيقة معيبة . والمطلوب إيجاد أن الرقيقة من إنتاج الآلة C ، أي إيجاد $P(C/M)$. باستخدام نظرية بيزر نجد أن :

$$\begin{aligned}
 P(C / M) &= \frac{P(C)P(M / C)}{P(A)P(M / A) + P(B)P(M / B) + P(C)P(M / C)} \\
 &= \frac{(0.10)(0.04)}{(0.60)(0.02) + (0.30)(0.30) + (0.10)(0.04)} \\
 &= \frac{4}{25} = 0.16
 \end{aligned}$$

مثال (21-5)

ثلاثة صناديق A , B , C ، في الصندوق الأول A ثلاثة كرات حمراء وخمسة كرات بيضاء ، في الصندوق الثاني B كرتان من اللون الأحمر وكرة بيضاء ، وفي الصندوق الثالث C كرتان من اللون الأحمر وثلاثة كرات من بيضاء . إذا اختير صندوق بطريقة عشوائية وسحبت منه كرة ووجد أن الكرة حمراء فما هو احتمال أن تكون من الصندوق الأول A .

الحل :

من معطيات السؤال نجد أننا نبحث عن احتمال الصندوق A بمعلومية أن الكرة حمراء . أي أن المطلوب إيجاد $P(A/R)$. من أجل إيجاد هذا الاحتمال أي $P(A/R)$ ، يجب علينا أولاً حساب $P(A \cap R)$ ، $P(R)$ ، أن احتمال أن يكون الصندوق الأول A قد اختير وأن تكون الكرة الحمراء قد سحبت منه يساوي :

$$\frac{1}{3} \times \frac{3}{8} = \frac{1}{8}$$

أي أن :

$$P(A \cap R) = 1/8 = 0.125$$

وحيث أنه توجد ثلاث طرق تؤدي إلى ظهور كرة حمراء فإن احتمال $P(R)$ يمكن الحصول عليه كما يلي :

$$\begin{aligned} P(R) &= \frac{1}{3} \times \frac{3}{8} + \frac{1}{3} \times \frac{3}{2} + \frac{1}{3} \times \frac{2}{5} \\ &= \frac{173}{360} \\ &= 0.48 \end{aligned}$$

ومنه نجد الاحتمال المطلوب وهو $P(A \cap R)$:

$$\begin{aligned} P(A / R) &= \frac{P(A \cap R)}{P(R)} \\ &= \frac{0.125}{0.48} \\ &= 0.26 \end{aligned}$$

كما يمكن إيجاد قيمة الاحتمال المطلوب أي أن يكون الصندوق الأول A قد اختير بمعلومية أن الكرة حمراء باستخدام نظرية بييز حيث أن :

$$\begin{aligned} P(A / R) &= \frac{P(A)P(R / A)}{P(A)P(R / A) + P(B)P(R / B) + P(C)P(R / C)} \\ &= \frac{(0.33)(0.375)}{(0.33)(0.375) + (0.33)(0.6) + (0.33)(0.4)} \\ &= 0.26 \end{aligned}$$

وهكذا حصلنا على نفس قيمة الاحتمال باستخدام نظرية بييز .

14.5 الأحداث المستقلة (Independent Events)

يقال لمجموعة من الأحداث بأنها مستقلة إذا كان وقوع أحدهما أو عدم وقوعه لا يؤثر في وقوع أي من باقي الأحداث . ومعنى هذا أنه يقال أن الحدث B مستقل عن الحدث A إذا كان احتمال حدوث B لا يتأثر بحدوث أو عدم حدوث A أو بمعنى آخر إذا كان احتمال B يساوي الاحتمال المشروط للحدث B عند وقوع الحدث A : $P(B/A)$. وبالتعويض عن قيمة $P(B/A)$ بالمقدار $P(B)$ في نظرية حاصل الضرب :

$$P(A \cap B) = P(A) P(B/A)$$

ومنه نحصل على :

$$P(A \cap B) = P(A) P(B)$$

وسوف نستخدم هذه العلاقة في التعريف الرياضي للاستقلال الحوادث حيث :

* يقال للحدثان A , B أنهما مستقلان إذا تحقق الشرط :

$$P(A \cap B) = P(A) P(B)$$

وإذا لم يتحقق هذا الشرط قيل أنهما غير مستقلين .

يتضح من التعريف مباشرة أن الحادثين A و B مستقلان إذا كان :

$$P(A/B) = P(B).P(A) \dots\dots\dots(12-5)$$

وذلك لأن :

$$P(A \cap B) = P(B).P(A/B) = P(B).P(A)$$

كما نلاحظ أن الشرطين :

$$P(B/A) = P(B) \quad \text{و} \quad P(A/B) = P(A)$$

متكافئان لأن :

$$P(A/B) = P(A).P(B/A) = P(B).P(A/B)$$

فأحدهما يقضي بحدوث الآخر أي عند التعويض عن قيمة أحدهما في المعادلة أعلاه سوف نجد قيمة الآخر .

مثال (22-5)

صندوق فيه 20 مصباح كهربائي 5 منها تالفة . سحبنا عشوائياً مصباحين على التوالي دون إعادة . ما هو احتمال أن يكون كلا المصباحين تالفين .

الحل :

نفرض أن الحادث A الذي يقع إذا وفقط إذا كان المصباح الأول تالفاً ، ومن الواضح أن :

$$P(A/B) = P(A)$$

لأنه لا علاقة بين تلف المصباح الأول وتلف المصباح الثاني إذن يتحقق الشرط :

$$P(A/B) = P(B).P(A)$$

وبما أن :

$$P(A) = \frac{5}{20} = \frac{1}{4} \quad , \quad P(B) = \frac{4}{19}$$

$$\begin{aligned} \therefore P(A/B) &= P(A) \cdot P(B) \\ &= \left(\frac{1}{4}\right) \left(\frac{4}{19}\right) = \frac{1}{19} \end{aligned}$$

لاحظ أنه لو كان المصباح بعد سحبه يعاد إلى الصندوق ثم يعاد إلى السحب ثانية بعد إعادة ترتيب المصابيح عشوائياً فإنه عندئذ نحصل على :

$$P(A) = \frac{1}{4} \quad , \quad P(B) = \frac{1}{4}$$

$$\begin{aligned} \therefore P(A/B) &= P(A) \cdot P(B) \\ &= \left(\frac{1}{4}\right) \left(\frac{1}{4}\right) = \frac{1}{16} \end{aligned}$$

مثال (5-23)

في تجربة إطلاق على هدف ما إذا كان احتمال أن يصيب الشخص الأول A يساوي 0.25 ، واحتمال أن يصيب الشخص الثاني B هو 0.4 . ما هو احتمال إصابة الهدف إذا صوب كل من A , B نحو الهدف مرة واحدة .

الحل :

من معطيات السؤال نجد أن لدينا :

$$P(A) = 0.25$$

$$P(B) = 0.40$$

ونلاحظ أيضا أن احتمال أن يصيب A أو احتمال أن يصيب B لا يتأثر بنتيجة الآخر ، وذلك يعني أن الحدث A يصيب الهدف مستقل عن الحدث B يصيب الهدف أي أن :

$$P(A \cap B) = P(A) P(B)$$

ومنه نجد أن :

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A) P(B) \\ &= 0.25 + 0.4 - (0.25 \times 0.4) \\ &= 0.55 \end{aligned}$$

وهكذا نجد أن احتمال إصابة الهدف إذا صوب كل من A و B نحو الهدف مرة واحدة يساوي 0.55 .

مثال (5-24)

إذا كان احتمال أن يعيش رجل 15 سنة أخرى هو 0.25 ، واحتمال أن تعيش زوجته 15 سنة أخرى هو 0.33 أوجد :

- 1- احتمال أن يعيش الزوج والزوجة 15 سنة أخرى .
- 2- احتمال أن يعيش احدهما على الأقل 15 سنة أخرى .
- 3- احتمال أن يموت الاثنان خلال الخمسة عشر سنة .
- 4- أن تعيش الزوجة 15 سنة .

الحل :

نفرض أن الحدث A هو أن يعيش الرجل 15 سنة ، والحدث B أن تعيش الزوجة 15 سنة وعليه فإن :

$$P(B) = 0.33 \quad , \quad P(A) = 0.25$$

1- لإيجاد احتمال أن يعيش الزوج والزوجة 15 سنة أخرى يجب علينا البحث عن $P(A \cap B)$ ، وحيث أن الحدثان مستقلان فإن :

$$\begin{aligned} P(A \cap B) &= P(A) P(B) \\ &= 0.25 \times 0.33 = 0.083 \end{aligned}$$

2- احتمال أن يعيش أحدهما على الأقل يعني أنه يجب البحث عن $P(A \cup B)$ حيث :

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.25 + 0.33 - 0.083 = 0.49 \end{aligned}$$

3- أما احتمال أن يموت الاثنان خلال 15 سنة فيعني أنه يجب إيجاد $P(A^c \cap B^c)$ حيث أن :

$$P(A^c) = 1 - P(A) = 1 - 0.25 = 0.75$$

$$P(B^c) = 1 - P(B) = 1 - 0.33 = 0.66$$

وحيث أن A^c , B^c حدثان مستقلان أيضاً فإن :

$$\begin{aligned} P(A^c \cap B^c) &= P(A^c) P(B^c) \\ &= 0.75 \times 0.66 = 0.49 \end{aligned}$$

4- أما احتمال أن تعيش الزوجة 15 سنة فيعني أنه يجب البحث عن $P(A^c \cap B)$ ، وحيث أن $P(A^c) = 0.75$ ، والحدثان A^c ، B مستقلان فإن :

$$\begin{aligned} P(A^c \cap B) &= P(A^c) P(B) \\ &= 0.75 \times 0.33 = 0.247 \end{aligned}$$

مثال (5-25)

في مسابقة رماية للمسدات على هدف متحرك إذا كان احتمال أن تصيب ثلاث سيدات هو 0.15 ، 0.25 ، 0.35 على التوالي . وكان كل منهم يصوب مرة واحدة على الهدف فأوجد :

- 1- احتمال أن تصيب سيدة واحدة منهم فقط الهدف .
- 2- احتمال أن تكون السيدة الأولى هي التي أصابت الهدف إذا أصيب الهدف من قبل سيدة واحدة فقط .

الحل :

نفرض أن الحدث C هو السيدة الأولى تصيب الهدف ، وأن الحدث D هو السيدة الثانية تصيب الهدف ، والحدث E هو السيدة الثالثة تصيب الهدف وعليه فإن :

$$P(C)=0.15 \quad , \quad P(D)=0.25 \quad , \quad P(E)=0.35$$

وهذه الأحداث الثلاثة مستقلة ، ونجد أيضا أن :

$$P(C^c)=0.85 \quad , \quad P(D^c)=0.75 \quad , \quad P(E^c)=0.65$$

أولاً : نفرض أن الحدث W هو حدث سيدة واحدة تصيب الهدف أنن :

$$W=(C \cap D^c \cap E^c) \cup (C^c \cap D \cap E^c) \cup (C^c \cap D^c \cap E)$$

أي أنه إذا كانت سيدة واحدة فقط قد أصابت الهدف فلا بد أن تكون الأولى فقط أي :

$$(C \cap D^c \cap E^c)$$

أو أن تكون السيدة الثانية فقط أي :

$$(C^c \cap D \cap E^c)$$

أو أن تكون السيدة الثالثة أي :

$$(C^c \cap D^c \cap E)$$

وبما أن هذه الأحداث الثلاثة متنافية فإننا نحصل على احتمال أن تصيب سيدة واحدة منهم الهدف هو :

$$\begin{aligned} P(W) &= P(C \cap D^c \cap E^c) \cup P(C^c \cap D \cap E^c) \cup P(C^c \cap D^c \cap E) \\ &= P(C)P(D^c)P(E^c) + P(C^c)P(D)P(E^c) + P(C^c)P(D^c)P(E) \\ &= (0.15)(0.75)(0.65) + (0.85)(0.25)(0.65) + (0.85)(0.75)(0.35) \\ &= 0.073 + 0.138 + 0.223 \\ &= 0.44 \end{aligned}$$

ثانياً : أما احتمال أن تكون السيدة الأولى هي التي أصابت الهدف فقط فإنه يعني أن علينا البحث عن $P(C/W)$ وهو احتمال أن تصيب السيدة الأولى الهدف بمعلومية أن سيدة واحدة قد أصابت الهدف ، وحيث أن :

$$C \cap W = C \cap D^c \cap E^c$$

هو حدث إصابة السيدة الأولى فقط للهدف ، فإن احتمال ذلك يساوي :

$$\begin{aligned} P(C \cap W) &= P(C \cap D^c \cap E^c) \\ &= 0.073 \end{aligned}$$

أما احتمال أن تصيب الهدف سيده فقط $P(W)$ فقد تم حسابه في المطلوب الأول ويساوي 0.44 ، وعليه نجد الاحتمال المشروط والمطلوب حيث :

$$P(C \cap W) = \frac{P(C \cap W)}{P(W)}$$

$$= \frac{0.073}{0.44} = 0.17$$

مثال (5-26)

تطلق طائرة صواريخ ضد أهداف معينة ، إذا كان احتمال إصابة الهدف لهذه الصواريخ يساوي 0.4 فما هو عدد الصواريخ التي يجب إطلاقها لكي يكون احتمال إصابة الهدف المطلوب على الأقل 90% .

الحل :

نفرض أن A هو حدث إصابة الهدف ، وعليه فإن $P(A) = 0.4$ ، ومنه فإن احتمال عدم إصابة الهدف هو $P(A^c) = 0.6$. لذا لإيجاد احتمال أن يخطأ عدد من الصواريخ n هو $(0.6)^n$ ، لذلك يجب علينا أن نبحث عن أصغر عدد صحيح n بحيث يكون :

$$1 - (0.6)^n > 0.9$$

أو أن :

$$(0.6)^n < 0.1$$

$$(0.6)^1 = 0.6 , (0.6)^2 = 0.36 , (0.6)^3 = 0.21 , (0.6)^4 = 0.13 , (0.6)^5 = 0.07$$

وهكذا يجب أن يكون عدد الصواريخ اللازم إطلاقها خمسة لكي يكون احتمال إصابة الهدف على الأقل 90% .

15.5 التكرار النسبي والاحتمالات التجريبية (Relative Frequency and Empirical Probabilities)

أشرنا سابقاً إلى أنه لوحظ تجريبياً أن النسبة $f = s/n$ ، والتي تسمى التكرار النسبي تصبح مستقرة في المدى الطويل أي أنها تقترب من نهاية ما ، ويعتبر هذا الاستقرار هو أساس نظرية الاحتمالات . وأن الأمثلة التالية سوف تساعدنا على فهم فكرة الاحتمالات التجريبية .

مثال (5-27)

لندرس بعض الافتراضات التالية :

- (a) تتكون أسرة معينة من أب وأم و 4 أطفال دون سن 16 سنة .
- (b) وتعيش هذه الأسرة في قرية فيها 400 طفل دون سن 16 سنة .
- (c) وتقع هذه القرية في محافظة بها 400,000 طفل دون سن 16 سنة .

أوجد نسبة الذكور من الأطفال في كل من a , b , c في كل من الافتراضات الثلاث الموجودة .

الحل :

يمكن أن نبني الحكم على مشاهداتنا السابقة ، حيث أنه لا بد أن كلاً منا شاهد أسرة كل أطفالها من الذكور وأخرى كل أطفالها من الإناث وغيرها وعليه فإن نسبة الذكور إما أن تكون صفراً أو 25 % أو 50% أو 75 % ، 100% ، وأخرى كل أطفالها من الإناث وغيرها .

في الافتراض الثاني لا نستطيع أن نصق بناءً على مشاهداتنا السابقة أن يكون جميع أطفال القرية ذكوراً أو جميعهم إناثاً بل لا نستطيع أن نصق أن نسبة الذكور من أطفال القرية هي 25 % أو 75 % ، حيث نتوقع أن تكون

نسبة الذكور واقعة بين 45 % - 55 % مثلاً من أطفال القرية ما لم تحدث هجرة للأطفال الذكور أو الأطفال الإناث .

وفي الافتراض الثالث إذا قيل لنا أن نسبة الذكور في المحافظة 54 % أو 46 % فإن ذلك سوف يكون غريباً نوعاً ما حيث أننا ننتظر أن تكون النسبة قريبة جداً من 50 % أو 51 % .

ومن دراستنا للإحصائيات السابقة ومن مشاهداتنا نعرف أنه كلما زاد عدد الأسر التي تأخذ بيانات عنها ، كلما اقتربت نسبة الذكور لجميع هؤلاء الأطفال من نسبة 50 % أو 51 % ، حيث تبدو هذه النسبة طبيعية وسوف لن تكون غريبة ، حتى ولو وجدنا فيها ابتعاداً كبيراً عن هذه القيمة بخصوص أسرة معينة .

مثال (5 - 28)

لندرس الافتراضات التالية :

- (a) يبلغ عمر رجل معين 40 عاماً .
 - (b) ويعيش هذا الرجل في قرية بها 75 رجلاً في نفس السن .
 - (c) تقع القرية في محافظة بها 75,000 في نفس السن .
- فكم ننتظر أن يكون عدد الأحياء من الرجال الموجودين بعد مضي سنة في كل من الافتراضات السابقة .

الحل :

في الافتراض الأول لا نستطيع أن نجزم بشيء فقد يعيش الرجل إلى العام القادم وقد يموت قبل ذلك . وفي الافتراض الثاني لن نستغرب إذا مات واحد مثلاً من الرجال خلال سنة . ولكننا لا نتوقع موت نسبة كبيرة منهم في سنة واحدة إلا إذا حدثت كارثة .

أما في الافتراض فإننا نميل إلى الاعتقاد بأن نسبة معينة من رجال المحافظة الذين عمرهم اليوم 40 سنة تماماً سيموتون خلال سنة . فلو كانت هذه النسبة مثلاً تساوي (0.006) ، لوجدنا أن عدد الوفيات يساوي :

$$75000 \times \frac{6}{1000} = 450$$

مثال (29-5)

في أحد المعاهد المهنية العليا قام طلبة ثلاثة فصول بإجراء تجربة ، حيث قام كل طالب برمي قطعة نقود فضية في الهواء وعندما استقرت القطع تم إحصاء عدد القطع التي كسرت بوجهها العلوي صورة . وقد كرر كل طالب هذه التجربة 10 مرات ، وهذا يكافئ رمي الطالب قطعة نقود 100 مرة ، وكانت النتيجة كما موضحة في الجدول (3-5) .

جدول (3-5)

نتائج مشاهدة لتجارب قطعة نقود

الفصل	عدد الطلبة	عدد الرميات	عدد الصور	نسبة الصور
الأول	232	232000	11821	50.9 %
الثاني	325	325000	16582	51.0 %
الثالث	152	152000	7450	49 %
ثلاثة فصول	709	709000	35853	50.6 %

فما الذي يمكننا أستنتاجه من هذه التجربة .

الحل :

من الواضح أنه إذا كان عدد رميات قطعة نقود من النوع الذي أجريت عليه التجربة كبيراً فإن ظهور الصورة في أعلى القطعة بعد استقرارها يميل إلى الحثوث في نصف الرميات تقريباً . وهذا ما يقودنا إلى التعريف التالي :

تعريف

إذا تعرض حدث للتجربة لعدد n من المرات وكانت نتيجة كل تجربة هي أن يقع الحادث أو لا يقع ، وكانت s هي عدد مرات وقوعه فإن النسبة $f = s/n$ تسمى التكرار النسبي (Relative Frequency) لوقوع الحدث في هذه المجموعة من التجارب أي أن :

$$\text{التكرار النسبي} = \frac{\text{عدد مرات وقوع الحدث}}{\text{عدد مرات تعرض الحدث للتجربة}} \quad (13-5)$$

وبين الجدول (5- 4) التكرار النسبي لظهور الصورة في تجربة رمي قطعة النقود وظهور الصورة .

جدول (5 - 4)

التكرار النسبي لظهور الصورة في عدد من محاولات إلقاء قطعة النقود

الفقرة	الفصل الأول	الفصل الثاني	الفصل الثالث	الفصول الثلاثة
عدد التجارب (n)	232000	325000	152000	709000
عدد مرات الوقوع (w)	11821	17450	16582	35853
التكرار النسبي ($\frac{w}{n}$)	0.509	0.510	0.490	0.506

وقد أدت هذه الاعتبارات والافتراضات أي استباق الحدث قبل وقوعه إلى ظهور ما يسمى بنظرية التكرار والتي تنص على أنه إذا تعرض حدث لتجربة لعدة مرات فإنه يمكن أخذ قيمة التكرار النسبي كتقدير لقيمة احتمال وقوع ذلك الحدث . إن الاحتمالات بهذا المعنى تدعى " الاحتمالات التجريبية " ونلاحظ أن القيم المستنتجة تساوي تقريباً الاحتمالات القبلية التي تم عرضها سابقاً .

16.5 التوقع الرياضي (Mathematical Expectation)

قبل البدء بدراسة التوقع الرياضي يجب علينا توضيح مفهوم المتغير العشوائي (Random Variable) . نفرض أن S فضاء عينة لتجربة ما ، وأردنا تخصيص عدد معين لكل ناتج ، مثل طول عمر مصباح كهربائي بالأيام ، مجموع العددين عند إلقاء حجر الزهر وغيرها ، حيث نلاحظ مما سبق دراسته أن نواتج التجربة أي نقط المعاينة في S لا تكون أعداداً دائماً .

ويسمى مثل هذا التخصيص بشكل عام بالمتغير العشوائي ، ويرمز له بالرمز X ويعرف على أنه دالة من S إلى مجموعة الأعداد الحقيقية R ، بحيث تكون الصورة العاكسة لأي فترة من مجموعة الأعداد الحقيقية R حدثاً في فضاء العينة S . ويجب الإشارة هنا إلى أنه إذا كان فضاء العينة S فضاءً متقطعاً حيث تعرف كل مجموعة جزئية حدثاً فإن كل دالة حقيقية على S هي متغير عشوائي .

فلو كان X هو متغير عشوائي معرف على فضاء العينة S بحيث تكون صورته مجموعة منتهية $X(S)$:

$$X(S) = \{ x_1, x_2, x_3, \dots, x_n \}$$

ووجدنا من $X (S)$ فضاء احتمال بتعريف احتمال x_i على أنه $P (X = x_i)$ ويكتب عادة على شكل دالة $f(x_i)$ ، وتسمى هذه الدالة f المعرفة على $X(S)$ $f(x_i)=P(X=x_i)$ بدالة التوزيع أو دالة الاحتمال المتغير X وتعطى عادة على صورة جدول وتحقق دالة التوزيع f الشروط التالية :

$$(1) \quad \sum_{i=1}^n f(x_i) = 1$$

$$(2) \quad f(x_i) \geq 0$$

وكان X متغيراً عشوائياً له دالة توزيع فإن التوقع أو القيمة المتوقعة للمتغير X والذي يرمز له بالرمز $E(X)$ يعرف حسب العلاقة التالية :

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \dots + x_n f(x_n) = \sum_{i=1}^n x_i f(x_i)$$

ويمكن أيضاً تعريف التوقع الرياضي $E(X)$ على أنه الوسط المرجح أو المقيم للقيم الممكنة للمتغير العشوائي ، حيث ترجح كل قيمة من المتغير باحتمالها . أن مفهوم ومعنى التوقع الرياضي أو القيمة المتوقعة ستتوضح أكثر من خلال حلول الأمثلة والتجارب التالية .

مثال (30-5)

توجد قطعة نقود معدنية مثقلة من أحد وجهيها بحيث إذا رميت مرات عديدة ظهرت الصورة في 0.3 من الرميات وظهرت الكتابة في 0.7 من الرميات . وقد أتفق A مع B على أن يرمي طفل قطعة النقود ويأخذ B من A مبلغ مقداره خمسون ديناراً إذا ظهرت الصورة ، ولا يأخذ شيئاً إذا ظهرت الكتابة فإذا أعيدت التجربة لعدة مرات فما هو متوسط المبلغ الذي يأخذه B في الرمية .

الحل :

إذا رميت قطعة النقود عدد n من المرات وكانت S كبيرة بدرجة كافية ، يكون عدد مرات ظهور الصورة يساوي $0.3S$. إذن مجموع ما يأخذه B يساوي $50 \times 0.3S$ ، وبالقسمة على S يكون متوسط ما يأخذه B في الرمية الواحدة هو :

$$\frac{0.3 S \times 50}{S} = 50 \times 0.3 = 15$$

ويسمى هذا المتوسط بالقيمة المتوقعة (Expected Value) وهو المبلغ الذي يأخذه A من B في الرمية . في هذا المثال نلاحظ ما يلي :

أولاً - للتجربة نتيجتان ممكنتان وهما :

ظهور الصورة واحتمال حدوث هذه النتيجة 0.3 أو ظهور الكتابة واحتمال حدوث النتيجة الأخرى 0.7 .

ثانياً - ربطنا المبلغ خمسون دينار بظهور الصورة ولم نربط شيئاً بظهور الكتابة .

ثالثاً - وجدنا أن القيمة المتوقعة للمبلغ الذي يأخذه B من A يساوي 0.3×50 أي يساوي المبلغ الذي ربطناه بظهور الصورة \times احتمال ظهور الصورة .

مثال (5-31)

في المثال السابق إذا اتفق كل من C و D على أن يرمي طفل قطعة النقود المذكورة ويأخذ C من D مبلغ خمسون ديناراً إذا ظهرت الصورة ، ومبلغ 20 ديناراً إذا ظهرت الكتابة . فإذا أعيدت هذه التجربة مرات عديدة فما هو متوسط المبلغ الذي يأخذه C من D في الرمية الواحدة .

الحل :

إذا رميت قطعة النقود عدد n من المرات وكانت n كبيرة يكون عدد مرات ظهور الصورة يساوي $0.3n$ وعدد مرات ظهور الكتابة يساوي $0.7n$ فإن مجموع ما يأخذه C يساوي :

$$50 \times 0.3n + 20 \times 0.7n$$

وبالقسمة على n يكون متوسط ما يأخذه C في الرمية يساوي :

$$0.7 \times 20 + 0.3 \times 50 = 14 + 15 = 29$$

إن هذا المتوسط هو ما نسميه القيمة المتوقعة وهو المبلغ الذي يأخذه C في الرمية وفي هذا المثال نلاحظ ما يلي :

أولاً - للتجربة نتيجتان ممكنتان :

ظهور الصورة واحتمال حدوث هذه النتيجة 0.3 أو ظهور الكتابة واحتمال حدوث النتيجة الأخرى 0.7 .

ثانياً - ربطنا المبلغ 50 دينار باحتمال ظهور الصورة وربطنا المبلغ 20 دينار باحتمال ظهور الكتابة .

ثالثاً - وجدنا أن القيمة المتوقعة للمبلغ الذي يأخذه C من D يساوي :

$$0.7 \times 20 + 0.3 \times 50$$

أي ما يعادل المبلغ الذي ربطناه باحتمال ظهور الصورة ضرب احتمال ظهور الصورة مضافاً إليه المبلغ الذي ربطناه باحتمال ظهور الكتابة ضرب احتمال ظهور الكتابة .

مثال (32-5)

أتفق شخصان A , B على رمي حجر نرد عادي بحيث يأخذ الشخص الأول A من الشخص الثاني B عدداً من الدنانير يساوي مربع العدد الذي يظهر على الوجه العلوي لحجر النرد . فإذا أعيدت هذه النتيجة لعدة مرات فكم هو متوسط المبلغ الذي يأخذه الشخص الأول من الشخص الثاني في الرمية .

الحل :

إذا أعيدت التجربة لعدد n من المرات وكانت n كبيرة بدرجة كافية فإننا نتوقع أن تكون عدد مرات ظهور العدد 1 يساوي :

$$\frac{1}{6}n$$

إذن مجموع ما يأخذه الشخص الأول A عند ظهور العدد 1 يساوي :

$$1^2 \times \frac{1}{6}n$$

وبالمثل مجموع ما يأخذه الشخص الأول A عند ظهور العدد 2 يساوي :

$$2^2 \times \frac{1}{6}n$$

إذن مجموع ما يأخذه الشخص الأول A في n من الرميات يعادل التالي :

$$1^2 \times \frac{1}{6}n + 2^2 \times \frac{1}{6}n + 3^2 \times \frac{1}{6}n + 4^2 \times \frac{1}{6}n + 5^2 \times \frac{1}{6}n + 6^2 \times \frac{1}{6}n$$

وبالقسمة على n نحصل على :

$$\begin{aligned}
& 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} \\
&= 1 \times \frac{1}{6} + 4 \times \frac{1}{6} + 9 \times \frac{1}{6} + 16 \times \frac{1}{6} + 25 \times \frac{1}{6} + 36 \times \frac{1}{6} \\
&= \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6} \\
&= \frac{1+4+9+16+25+36}{6} \\
&= \frac{91}{6} = 15\frac{1}{6}
\end{aligned}$$

وهذا المتوسط يمثل المبلغ الذي يأخذه الشخص الأول A من الشخص الثاني B . وفي هذا المثال نلاحظ ما يلي :

أولاً - للتجربة 6 نتائج محتملة الظهور وهي العدد 1 أو العدد 2 , , , , , أو العدد 6 واحتمال ظهور أي من هذه النتائج يساوي $\frac{1}{6}$.

ثانياً - ربطنا بظهور أي عدد مبلغاً من القروش يساوي مربع هذا العدد .

ثالثاً - وجدنا أن القيمة المتوقعة للمبلغ الذي يأخذه الشخص الأول A من الشخص الثاني B يساوي القيمة :

$$1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6}$$

أي يساوي المبلغ الذي ربطناه بظهور العدد 1 ضرب احتمال ظهور العدد 1 مضافاً إليه المبلغ الذي ربطناه بظهور العدد 2 ضرب احتمال ظهور العدد 2 + المبلغ الذي ربطناه بظهور العدد 6 ضرب احتمال ظهور العدد 6 .

مثال (5-33)

في تجربة سحب ورقة بورقة من مجموعة أوراق الكوتشينة بدون إعادة الورقة المسحوبة حتى تظهر صورة الولد السباتي . أوجد متوسط عدد الأوراق التي تسحب حتى تظهر تلك الصورة .

الحل :

نفرض تساوي فرص السحب لكل أوراق المجموعة . حيث أنه يوجد 52 ورقة في مجموعة أوراق اللعب ، فإن احتمال ظهور الصورة المسحوبة في أي سحب معين يساوي $\frac{1}{52}$ ومعنى هذا إذا اعتبرنا أن سحب ورقة بورقة من مجموعة أوراق اللعب لغاية ظهور الولد السباتي تجربة ، فإنه إذا أجريت التجربة لعدد مساوي لـ n من المرات فإننا نتوقع ما يلي :

ظهور الصورة المطلوبة في مرة واحدة $S \times \frac{1}{52}$ من التجارب أو ظهور الصورة المطلوبة بعد :

$$2 \times \frac{1}{52} S$$

أو بعد :

$$3 \times \frac{1}{52} S$$

وهكذا لغاية الوصول إلى :

$$52 \times \frac{1}{52} S$$

فيكون المجموع يساوي :

$$1 \times \frac{1}{52} S + 2 \times \frac{1}{52} S + 3 \times \frac{1}{52} S + \dots + 52 \times \frac{1}{52} S$$

وبالقسمة على S يكون متوسط الأوراق المسحوبة في التجربة لغاية ظهور صورة الولد السباتي مساوي إلى :

$$\begin{aligned} & 1 \times \frac{1}{52} + 2 \times \frac{1}{52} + 3 \times \frac{1}{52} + 4 \times \frac{1}{52} + \dots + 52 \times \frac{1}{52} \\ &= \frac{1}{52} + \frac{2}{52} + \frac{3}{52} + \frac{4}{52} + \frac{5}{52} + \dots + \frac{52}{52} \\ &= \frac{1378}{52} = 26 \frac{26}{52} = 26.5 \end{aligned}$$

17.5 قاعدة حساب القيمة المتوقعة

(Expected Value Approximation Rule)

من البند السابق نجد أنه إذا ربطت قيمة ما n بحدث A احتمال وقوعه P فإن القيمة المتوقعة للحدث يساوي الاحتمال $x.P(A)$ عدد مرات الوقوع n ، وقد استخدم نمط التوقع (Expectation) بدلاً من القيمة المتوقعة .

وعليه إذا ربطت القيم $n_1, n_2, n_3, \dots, n_m$ بالأحداث $A_1, A_2, A_3, \dots, A_m$ فإن التوقع الرياضي أو القيمة المتوقعة تعطى حسب العلاقة التالية :

$$E = P_1(A).n_1 + P_2(A).n_2 + \dots + P_m(A).n_m \dots \dots \dots (14-5)$$

وهكذا نجد من هذه العلاقة أن التوقع الرياضي E هو الوسط المرجح أو المقيم للقيم الممكنة للمتغير العشوائي ، حيث ترجح كل قيمة من المتغير باحتمالها كما تم الإشارة إليه سابقاً .

مثال (5-34)

إذا أُلقي حجر نرد مرة واحدة . فأوجد احتمال ظهور عدد أقل أو يساوي 4 واحتمال ظهور عدد أكبر من 4 .

الحل :

الحدث أقل أو يساوي 4 هو الحدث :

$$\{1,2,3,4\}$$

ويتكون من أربعة عناصر من العناصر الستة لفضاء العينة أثناء إلقاء حجر النرد والحدث أكبر من 4 هو الحدث :

$$\{5,6\}$$

إذن :

$$P(n > 4) = \frac{2}{6} \quad \text{و} \quad P(n \leq 4) = \frac{4}{6}$$

لاحظ أن مجموع الاحتمالين يساوي 6 لأن الحدث الثاني لا يقع إذا ما وقع الحدث الأول ونعبر عن هذا بأن نقول أن الحدثين هما شاملين ومانعين بالتبادل .

وعلى وجه العموم إذا كانت نتائج تجربة صنفه تحتويها الأحداث الشاملة المانعة بالتبادل $A_1, A_2, A_3, \dots, A_m$ فإن :

$$P(A_1) + P(A_2) + P(A_3) + \dots + P(A_m) = 1 \quad \dots\dots(15-5)$$

وصيغة الشمول هنا معناها أنه لا يمكن أن يقع أي حادث آخر نتيجة التجربة .

مثال (35-5)

إذا ألقى حجراً مرة واحدة وكان P_R هو الحادث الذي يقع إذا كان المجموع على الوجهين العلويين مساوياً لـ R فالمطلوب بيان العناصر التي يتكون منها كل من الأحداث $P_1, P_2, P_3, \dots, P_{21}$ ثم إيجاد قيمة كل من :

- احتمال الحصول على مجموع زوجي .
- احتمال الحصول على مجموع يقبل القسمة على 3 .
- احتمال الحصول على مجموع زوجي ويقبل القسمة على 3 .
- احتمال الحصول على مجموع زوجي أو يقبل القسمة على 3 .
- احتمال أن يكون المجموع يقبل القسمة على 3 إذا علم أنه زوجي .

الحل :

نفرض أن (x, y) رمزاً للحادث حيث أن العدد على الحجر الأول x ، والعدد على الحجر الثاني y .

$\sum R$ - رمزاً للحادث ، المجموع على الحجرين يساوي r .

A - رمزاً للحادث ، المجموع عند زوجي .

B - رمزاً للحادث ، المجموع يقبل القسمة على 3 .

$P(R)$ - رمزاً لاحتمال الحصول على مجموع يساوي R .

أن الجدول (5-5) يحتوي على جميع النتائج الممكنة لإلقاء حجرى النرد معاً وقد جمعت النتائج التي تكون الحدث (P_R) معاً حيث :

$$(R = 1, 2, 3, \dots, 12)$$

فمثلاً الحدث P_2 يتكون من النتيجة الوحيدة (1,1) بينما الحادث P_3 سوف يتكون من النتيجتين (2,1) , (1,2) وهكذا .

(a) احتمال الحصول على مجموع زوجي هو :

$$\begin{aligned} P(A) &= P(2) + P(4) + P(6) + P(8) + P(10) + P(12) \\ &= \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{5}{36} + \frac{3}{36} + \frac{1}{36} = \frac{18}{36} = \frac{1}{2} \end{aligned}$$

(b) احتمال الحصول على مجموع يقبل القسمة على 3 هو :

$$\begin{aligned} P(B) &= P(3) + P(6) + P(9) + P(12) \\ &= \frac{2}{36} + \frac{5}{36} + \frac{4}{36} + \frac{1}{36} = \frac{12}{36} = \frac{1}{3} \end{aligned}$$

(c) احتمال الحصول على مجموع زوجي ويقبل القسمة على 3 هو :

$$\begin{aligned} P(A \cap B) &= P(6) + P(12) \\ &= \frac{5}{36} + \frac{1}{36} = \frac{6}{36} = \frac{1}{6} \end{aligned}$$

جدول (5-5)

الحوادث الشاملة المانعة بالتبادل التي يحتمل وقوعها إذا القى حجرا نرد لمرة واحدة فقط

P_R	R	P_R	R
$(1,2); (2,1)$	3	$(1,1)$	2
$(1,4); (2,3); (3,2); (4,1)$	5	$(1,3); (2,2); (3,1)$	4
$(4,3); (5,2); (6,1)$ $(1,6); (2,5); (3,4)$	7	$(5,1); (4,2); (5,1)$ $(1,5); (2,4); (3,3)$	6
$(3,6); (4,5); (5,4); (6,3)$	9	$(4,4); (5,3); (6,2)$ $(2,6); (3,5)$	8
$(5,6); (6,5)$	11	$(4,6); (5,5); (6,4)$	10
		$(6,6)$	12

(d) احتمال الحصول على مجموع زوجي أو يقبل القسمة على 3 هو :

$$\begin{aligned}
 P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
 &= \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} \\
 &= \frac{3+2-1}{6} = \frac{4}{6} = \frac{2}{3}
 \end{aligned}$$

(e) احتمال الحصول على مجموع يقبل القسمة على 3 إذا علم أنه زوجي هو :

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{1/6}{1/2} = \frac{2}{6} = \frac{1}{3}$$

ويجب ملاحظة أنه يمكن إيجاد الإجابات السابقة من الجدول (4-5) مباشرة ، فمثلاً تقع جميع الأحداث التي يكون فيها المجموع عدداً زوجياً في النصف الأيمن من الجدول وعدد نتائجها معاً 18 ، وبذلك يكون الاحتمال المطلوب في البند (a) يساوي $\frac{18}{36}$.

والنتائج التي تعطي مجموعاً يقبل القسمة على 3 هي نتائج الحوادث P_3 ، P_6 ، P_{12} و P_9 وعددها :

$$1 + 4 + 5 + 3 = 12$$

و يكون الاحتمال المطلوب في البند (b) في أعلاه هو $\frac{12}{36}$. ونترك للطالب إيجاد بقية الاحتمالات من الجدول .

مثال (5-36)

اتفق شخصان على أن يرمي الأول قطعة نقود 3 مرات بطريقة عشوائية ، ويأخذ من الثاني عدداً من الدنانير يساوي مربع عدد الصور التي تظهر . أوجد القيمة المتوقعة للمبلغ الذي يأخذه الأول من الثاني .

الحل :

لو رمزنا لظهور الصورة (Head) بالرمز H والكتابة (Tail) بالرمز T فإن إحدى النتائج الممكنة هي (H , H , T) أي ظهور الصورة في المرة

الأولى وظهورها في المرة الثانية وظهور الكتابة في المرة الثالثة . ويحتوي الجدول (5-6) على جميع النتائج الممكنة لتجربة إلقاء قطعة نقود 3 مرات متتالية .

جدول (5-6)

النتائج الممكنة لإلقاء قطعة نقود ثلاث مرات متتالية حيث $x =$ عدد مرات ظهور الصورة في الرميات الثلاثة

النتائج الممكنة أو المحتملة	x
(T, T, T)	0
(T, T, H); (T, H, T); (H, T, T)	1
(T, H, H); (H, T, H); (H, H, T)	2
(H, H, H)	3

وبلاحظ أن عدد النتائج الممكنة كلها 8 ، فإذا اعتبرنا أن كل هذه النتائج متساوية الإمكان فإن احتمال حدوث أي نتيجة معينة منها يساوي $1/8$. والسؤال هو ما هو احتمال الحصول على صورة واحدة بالضبط في الرميات الثلاثة بغض النظر عن موضع الرمية التي ظهرت فيها . للإجابة على هذا السؤال نلاحظ أن الحدث صورة واحدة في الرميات الثلاثة يمكن أن يتم بإحدى ثلاث نتائج مانعة بالتبادل كما يبين الجدول عند $x = 1$ واحتمال كل من هذه النتائج يساوي $1/8$.

فإذا كتبنا $P(x)$ يساوي احتمال الحصول على x بالضبط من الصور في الرميات الثلاث فإن احتمال الحصول على صورة واحدة هو :

$$P(1) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

وبنفس الطريقة يمكن إكمال الجدول (7-5) .

جدول (7-5)

احتمال الحصول على كل عدد ممكن من الصور عند إلقاء قطعة نقود 3 مرات متتالية

3	2	1	0	X
$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	P(X)

نلاحظ أن مجموع الاحتمالات يساوي 1 . إذن القيمة المتوقعة المطلوبة في المسألة تحسب بالشكل التالي :

$$\begin{aligned}
 & (0)^2 \times P(0) + (1)^2 \times P(1) + (2)^2 \times P(2) + (3)^2 \times P(3) \\
 &= (0) \frac{1}{8} + (1) \frac{3}{8} + (4) \frac{3}{8} + (9) \frac{1}{8} \\
 &= \frac{3}{8} + \frac{12}{8} + \frac{9}{8} = \frac{24}{8} = 3
 \end{aligned}$$

أن هذا الجواب يعني أنه إذا تكررت التجربة مرات عديدة بدرجة كافية فإن متوسط ما يأخذه A من B بعد إلقاء قطعة النقود ثلاث مرات يساوي تقريباً 3 قروش .

يحتوي صندوق على 4 كرات صفراء و 6 زرقاء ويحتوي صندوق آخر على 3 كرات صفراء و 5 زرقاء . وكل الكرات متشابهة فيما عدا اللون . وقد طلب من طفل أن يسحب كرة عشوائياً من كل صندوق . واتفق شخص مع شخص آخر على أن يدفع الأول للثاني مبلغ 20 دينار إذا كانت الكرتان المسحوبتان من لون واحد وأن يدفع 40 دينار إذا كانت الكرتان من لونين مختلفين . أوجد القيمة المتوقعة للمبلغ الذي يدفعه الشخص الأول .

الحل :

نفرض أن $P(Y)$ رمزاً لاحتمال سحب كرة صفراء من الصندوق الأول ونفرض $P(B)$ لاحتمال سحب كرة زرقاء منه . وسنكتب $P'(Y)$ و $P'(B)$ للاحتمالين المناظرين للصندوق الثاني فيكون :

$$P(Y) = \frac{4}{10}, P(B) = \frac{6}{10}$$

$$P'(Y) = \frac{3}{8}, P'(B) = \frac{5}{8}$$

وبما أن نتيجة السحب من الصندوق الثاني لا تتوقف على نتيجة السحب من الصندوق الأول فإن احتمال سحب كرة صفراء من الصندوق الأول وكرة صفراء من الصندوق الثاني هو :

$$P(Y) \cdot P'(Y) = \frac{4}{10} \times \frac{3}{8} = \frac{12}{80} = \frac{6}{40}$$

احتمال سحب كرة زرقاء من الكيس الأول وكرة زرقاء من الكيس الثاني هو :

$$P(B).P'(B) = \frac{6}{10} \times \frac{5}{8} = \frac{30}{80} = \frac{15}{40}$$

فإذا رمزنا للحدث سحب كرتين صفراوين بالرمز A والحدث سحب كرتين زرقاوين بالرمز B فيكون الحدث كرتان من لون واحد هو الحدث $A \vee B$ (A أو B) مع ملاحظة أن A ، B مانعان بالتبادل . إذن احتمال الحصول على كرتين من لون واحد هو :

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) \\ &= \frac{6}{40} + \frac{15}{40} \\ &= \frac{21}{40} = \frac{1}{6} \end{aligned}$$

وا احتمال الحصول على كرتين من لونين مختلفين يساوي :

$$\frac{19}{40} = \frac{21}{40} - 1$$

إذن القيمة المتوقعة للمبلغ الذي يدفعه الأول يساوي احتمال سحب كرتين من لون واحد ضرب 20 مضافاً إليه احتمال سحب كرتين من لونين مختلفين ضرب 40 أي يساوي :

$$20 \times \frac{21}{40} + 40 \times \frac{19}{40} = \frac{21}{2} + 19 = 10.5 + 19 = 29.5$$

مع ملاحظة احتمال أن تكون الكرتان من لونين مختلفين .

القيمة المتوقعة للمبلغ يساوي احتمال أن تكون الكرة الأولى صفراء والثانية زرقاء مضافاً إليه احتمال أن تكون الكرة الأولى زرقاء والثانية صفراء أو ما يساوي :

$$\frac{4}{10} \times \frac{5}{8} + \frac{6}{10} \times \frac{3}{8} = \frac{10}{40} + \frac{9}{40} = \frac{19}{40}$$

وهي نفس القيمة المحسوبة للاحتمال بالاعتماد على احتمال سحب كرتين من لون واحد .

18.5 التحليل التوافقي (The Harmonic Analysis)

إن إيجاد عدد عناصر فضاء العينة قد يكون سهلاً في بعض الأحيان وقد يصبح صعباً في أحيان أخرى . إذ أن فضاء العينة قد يحتوي على عدداً كبيراً من العناصر . لذا سنقوم في البنود القادمة بعرض بعض القواعد الرياضية المساعدة والمفيدة في حساب عدد عناصر فضاء العينة ، وبالتالي في حساب احتمال حدث ما .

1.18.5 المبدأ الأساسي في العد

(Fundamental Principles of Counting)

في هذا البند سوف نقوم بعرض بعض طرق تحديد عدد النواتج الممكنة لتجربة معينة أو عدد العناصر في مجموعة معينة بغير طريقة العد المباشرة المعروفة . وهذه الطرق تسمى كما أشرنا سابقاً بالتحليل التوافقي . ويمكن التعبير عن المبدأ الأساسي في العد بالشكل التالي :

إذا كان لدينا A_1, A_2, \dots, A_k فئات مختلفة عدد عناصرها على الترتيب هو n_1, n_2, \dots, n_k ، فإن عدد الطرق الممكنة لاختيار عنصر واحد من كل منها هو :

$$n_1.n_2.n_3.KK n_k = n! \dots\dots\dots (16 - 5)$$

إن الرمز $n!$ يسمى بمضروب العدد n ، ويظهر في أحيان كثيرة في الرياضيات عند إيجاد حاصل ضرب الأعداد الصحيحة الموجبة من العدد 1 حتى العدد n ويعرف على النحو التالي :

$$n! = 1.2.3.4.5.....(n - 2)(n - 1)n$$

ومن المفيد أن نعرف أن $0! = 1$.

19.5 التباديل والترتيبات والتوافيق

(Permutations , Arrangements and Combinations)

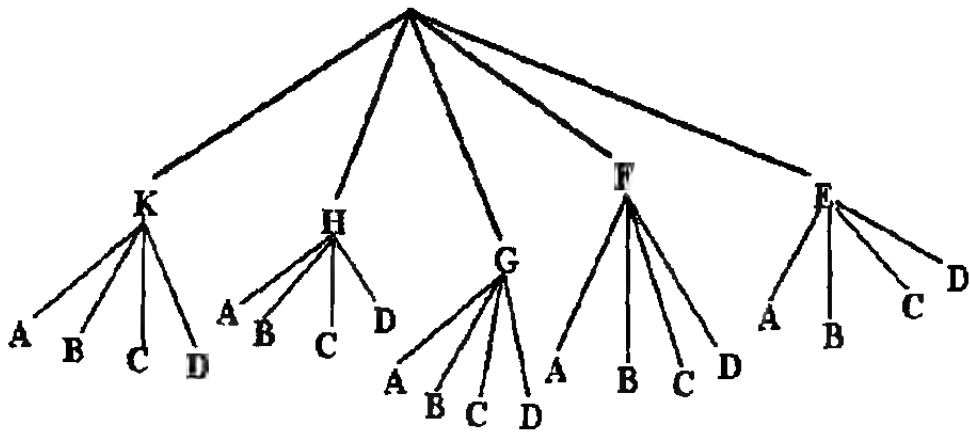
تستخدم هذه الطرق في تحديد عدد النواتج الممكنة لتجربة معينة أو عدد العناصر في مجموعة معينة بغير طريقة العد المباشر ، وتوضح من خلال دراسة الأمثلة التالية :

مثال (5-38)

أعلنت شركة تجارية عنوظيفتين شاغرتين فتقدم 5 رجال و4 سيدات لملئ هاتين الوظيفتين . أوجد بكم طريقة يمكن تعيين رجل وامرأة في هاتين الوظيفتين .

الحل :

نفرض أن النساء هن A , B , C و الرجال هم E , F , G , H و K وعليه فإن اختيار الشجرة البيانية (Graphical Tree) يكون على النحو الموضح في الشكل (5-2) .



الشكل (5 - 2)

الشجرة البيانية للاحتتمالات المتوقعة لتعيين رجل وامرأة لوظيفتين شاغرتين

نلاحظ أن هناك 20 طريقة مختلفة لاختيار رجل وامرأة من المرشحين التسعة .

1.19.5 التباديل (Permutations)

يسمى وضع n من الأشياء في ترتيب معين بأنه تبديل لهذه الأشياء بشرط أن تؤخذ جميع هذه الأشياء ، ويسمى وضع a عدد مثل r حيث r أقل أو يساوي n من هذه الأشياء في ترتيب معين بأنه تبديل العدد n من الأشياء مأخوذة r في كل مرة ، وسيرمز لعدد التباديل من الأشياء المأخوذة r كل مرة بالرمز :

$$P(n, r)$$

مثال (5-39)

ليكن لدينا ثلاثة عناصر هي a, b, c يمكن تشكيل ستة ثلاثيات مرتبة من هذه العناصر مع ملاحظة عدم تكرار العنصر في الثلاثية الواحدة . كم هو عدد تباديل هذه العناصر .

الحل :

العناصر الثلاثة ترتب بالشكل التالي :

$$(a, b, c), (a, c, b), (b, a, c), (b, c, a), (c, a, b), (c, b, a)$$

إن هذه العناصر تبادلت المواقع فيما بينها فكان عدد الأوضاع المختلفة الممكنة هو 6 أي أن عدد التباديل يساوي 6 .

مثال (5-40)

قطار مكون من 10 عربات وقاطرة . أوجد بكم طريقة يمكن ترتيب هذه العربات خلف القاطرة .

الحل :

$$P_{10} = 10! = (10)(9)(8)(7)(6)(5)(4)(3)(2)(1) = 3628800$$

وهكذا نجد أنه يمكن ترتيب العشر عربات خلف القاطرة بـ 3628800 طريقة .

نتيجة : عدد تباديل n من الأشياء مأخوذة جميعها بنفس الوقت هو $n!$.

مثال (5-41)

ما هو عدد تباديل ثلاثة عناصر مثل a, b, c مأخوذة جميعها بنفس الوقت .

الحل :

من النتيجة السابقة نجد أنه يوجد لدينا ثلاث تباديل وعليه فإن :

$$3! = 1 \cdot 2 \cdot 3 = 6$$

وهذه التباديل هي :

$$abc, acb, bac, bca, cab, cba$$

ويمكننا الآن اشتقاق الصيغة العامة للمقدار $P(n, r)$ بنفس الطريقة المتبعة في الأمثلة السابقة حيث يمكننا اختيار العنصر الأول في تبديل عناصر عددها n مأخوذة r كل مرة بطرق مختلفة عددها n ، وبعد ذلك يمكن اختيار العنصر الثاني في هذا التبديل بطرق عددها $n-1$ ، وبالاستمرار على هذا النحو يمكن اختيار العنصر الأخير الذي يكون ترتيبه r بطرق عددها هو :

$$n-(r-1) = n-r+1$$

ومنه نجد أن :

$$P(n, r) = \frac{n!}{(n-r)!} \dots\dots\dots (17-5)$$

وفي بعض الحالات الخاصة $r = n$ نجد أن :

$$P(n, n) = n(n-1)(n-2) \dots\dots\dots 4.3.2.1 = n!$$

2.19.5 الترتيب (Arrangements)

يعرف الرمز $\binom{n}{r}$ لعددين صحيحان موجبان n, r بحيث يكون $r \leq n$ كما يلي :

$$\binom{n}{r} = \frac{n(n-1)(n-2) \dots\dots\dots (n-r+1)}{1.2.3.4 \dots\dots\dots (r-1)r}$$

وتسمى هذه الأعداد بمعاملات ذات الحدين . كما سيتم توضيحه عند تعريف نظرية ذات الحدين لاحقاً .

مثال (5-42)

أوجد قيمة المعامل $\binom{10}{4}$.

الحل :

$$\binom{10}{4} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = 210$$

ويمكن الحصول من العلاقة والتعريف السابق على صيغة هامة من خلال

ملاحظة أنه يوجد r معامل في كل من البسط والمقام للعدد $\binom{n}{r}$ ، حيث :

$$\begin{aligned} \binom{n}{r} &= \frac{n(n-1)(n-2)\dots(n-r+1)}{1 \cdot 2 \cdot 3 \cdot 4 \dots (r-1)r} = \frac{n(n-1)\dots(n-r+1)(n-r)!}{1 \cdot 2 \cdot 3 \cdot 4 \dots (r-1)r(n-r)!} \\ &= \frac{n!}{r!(n-r)!} \end{aligned}$$

وبفضل هذه العلاقة أو الطريقة الثانية نستطيع توفير الوقت والعمليات الحسابية .

مثال (5-43)

أوجد قيمة المعامل $\binom{10}{7}$.

الحل :

باستخدام الطريقة الثانية نجد أن :

$$\binom{10}{7} = \binom{10}{3} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120$$

بالمقارنة مع الصيغة الأولى حيث :

$$\binom{10}{7} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 120$$

مثال (44-5)

ليكن لدينا 4 عناصر هي a, b, c, d . أوجد كم ثنائية مرتبة يمكن تشكيلها من هذه العناصر مع عدم تكرار العنصر الواحد في الثنائية الواحدة .

الحل :

إن عدد هذه الثنائيات هو 12 وهي كما يلي :

$$(a, b), (c, a), (c, d), (b, a), (b, c), (b, d) \\ (c, b), (a, c), (a, d), (d, a), (d, b), (d, c)$$

مثال (45-5)

إذا كان عدد المهندسين المرشحين للعمل في مصنع هو 12 مهندس وذلك لشغل أربعة وظائف مختلفة . أوجد بكم طريقة يمكن ملء هذه الوظائف الأربعة .

الحل :

$$A_n^k = A(n, k) = \frac{12!}{4!} \\ = 19958400$$

3.19.5 التوافيق (Combinations)

نفرض أن لدينا تجمعاً من n من الأشياء . يعرف توافق (Combinations) هذه الأشياء والتي عددها n مأخوذة r كل مرة (r - توافق) بأنه أي مجموعة جزئية بها r من الأشياء . وبمعنى آخر فإن التوافق r هو أي اختيار لعدد r شيء من بين هذه الأشياء ودون اعتبار لطريقة الترتيب .

فإذا رمزنا بالرمز $\binom{n}{r}$ أو بالرمز C_r^n أو بالرمز $C(n, r)$ لعدد توافق r عنصر مأخوذة معاً من n عنصر فإن :

$$C(n, r) = \binom{n}{r} = \frac{P(n, r)}{r!} = \frac{n!}{r!(n-r)!} \dots\dots (19-5)$$

وأيضاً فإن:

$$A_r^n = n! C_r^n \dots\dots\dots (20-5)$$

ويجب الإشارة إلى أن معاملات ذات الحدين $\binom{n}{r}$ قد عرفت بأنها :

$$\frac{n!}{r!(n-r)!}$$

وسوف نستمر باستخدام $C(n, r)$ و $\binom{n}{r}$ بنفس المعنى .

مثال (46-5)

لنكن لدينا العناصر a, b, c, d . أوجد بكم طريقة يمكن تشكيل فئات تضم عنصرين فقط منها .

الحل :

إن الفئات المشكلة هي :

$$\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$$

وعدد هذه الفئات يساوي 6 .

مثال (47-5)

أوجد بكم طريقة يمكن اختيار أربعة أوراق من أوراق الكوتشينية الشدة نو الاثنان وخمسون ورقة .

الحل :

عدد الطرق الممكنة لاختيار أربعة أوراق يمكن الحصول عليها باستخدام العلاقة (19-5) حيث :

$$C_{52}^4 = \frac{52!}{(4!)(48!)} = \frac{(52)(51)(50)(49)(48!)}{(4)(3)(2)(1)(48!)} = 270725$$

وهكذا نجد أنه يمكن اختيار أربعة أوراق من أوراق الكوتشينية الشدة بـ 270725 طريقة .

مثال (5-47)

أوجد كم لجنة رباعية يمكن تكوينها من ثمانية أشخاص .

الحل :

أن كل لجنة تعتبر توافق من الثمانية أشخاص مأخوذة منهم ثلاثة كل مرة
إذا يوجد :

$$C(8, 4) = \binom{8}{4} = \frac{8.7.6.5}{4.3.2.1} = \frac{1680}{24} = 70$$

وهكذا بسبعين طريقة يمكن تشكيل لجنة رباعية مختلفة من ثمانية أشخاص .

مثال (5-48)

في أحد المصانع يختار كل عام وفد من أربعة مهندسين عاملين في هذا
المصنع لتمثيل المصنع في معرض المنتجات الدولي الذي يقام سنوياً في بلد ما
أوجد :

(1) عدد الطرق التي يمكن اختيار الوفد بها إذا كان عدد المهندسين المرشحين
لذلك وتتوفر بهم الشروط هو 16 مهندس .

(2) عدد الطرق التي يمكن الاختيار إذا كان اثنان من المهندسين المرشحين لا
يستطيعان حضور المعرض في نفس الوقت أو بمعنى آخر معاً .

الحل :

أولاً : يمكن اختيار المهندسين الأربعة عن طريقة إيجاد توافق أي عن
طريق إيجاد :

$$\binom{16}{4} = \frac{16.15.14.13}{4.3.2.1} = 1820$$

ثانياً : نفرض أن المهندسين اللذان لا يمكن اختيارهما معاً هما D , C فيكون هناك طريقتان للحل ، الطريقة الأولى هي أنه إذا لم يتكون الوفد من C ولا من D . وفي هذه الحالة يمكن اختيار الوفد كما يلي :

$$({}^{14}_4) = \frac{14 \cdot 13 \cdot 12 \cdot 11}{4 \cdot 3 \cdot 2 \cdot 1} = 1001$$

أما إذا تكون الوفد من C أو من D ولكن ليس منهما معاً فأن الوفد يمكن اختياره بطرق عددها :

$$2 \cdot ({}^{14}_3) = 2 \cdot \frac{14 \cdot 13 \cdot 12}{3 \cdot 2 \cdot 1} = 728$$

وبذلك تكون الطرق الكلية لاختيار الوفد هي :

$$1001 + 728 = 1729$$

أما الحل بالطريقة الثانية فيتضمن أنه إذا تكون الوفد من المهندسين D و C معاً فأن المهندسين الآخرين يمكن اختيارهما بطرق عددها :

$$({}^{14}_2) = \frac{14 \cdot 13}{2 \cdot 1} = 91$$

وبذلك يكون عدد الطرق التي يمكن اختيار الوفد بها إذا لم يتكون الوفد من D , C معاً هو :

$$1820 - 91 = 1729$$

وهكذا حصلنا على نفس عدد الطرق والتي تساوي 1729 وهي النتيجة نفسها . ولكن الريق الثانية في الحل تعتبر أبسط وأسرع .

مثال (5-48)

في إحدى الشركات الكبرى إذا كان امتحان القبول لشغل وظيفة ما يتكون من عشرة أسئلة ، وكان يجب على المتقدم لشغل هذه الوظيفة أن يجيب على سبعة أسئلة فقط أوجد :

- (1) عدد الطرق التي يجب على المتقدم أن يختار بها الاسئلة .
- (2) عدد الطرق التي يمكنه الاختيار إذا كانت الاسئلة الاربعة الاولى إجبارية .
- (3) عدد الطرق التي يمكنه الاختيار إذا كان من الضروري أن يجيب على أربعة أسئلة من الاسئلة الخمسة الاولى .

الحل :

أولاً : يمكن للمتقدم اختيار الاسئلة السبعة بطرق عددها كما يلي :

$$\binom{10}{7} = \binom{10}{3} = \frac{10.9.8}{3.2.1} = 120$$

ثانياً : إذا قام المتقدم بالاجابة على الاسئلة الاربعة الاولى ، فيمكنه بعد ذلك أن يختار الاسئلة الثلاثة الاخرى من بين الاسئلة الستة الاخيرة بطرق عددها :

$$\binom{6}{3} = \frac{6.5.4}{3.2.1} = 20$$

ثالثاً : إذا أجاب المتقدم على الاسئلة الخمسة الاولى فيمكنه اختيار السؤاليين الآخرين من بين الخمسة أسئلة الاخيرة بطرق عددها :

$$\binom{5}{2} = \frac{5.4}{2.1} = 10$$

ومن ناحية أخرى إذا أجاب المتقدم على أربعة أسئلة فقط من بين الأسئلة الخمسة الأولى فيمكنه اختيار الأسئلة الثلاثة الأخرى من بين الأسئلة الخمسة الأخيرة بطرق عددها :

$$\binom{5}{3} = \frac{5.4.3}{3.2.1} = 10$$

وبذلك يمكنه اختيار الأسئلة السبعة بطرق عددها $100 = 10 \cdot 10$ ويكون عدد الاختيارات الكلية 105 اختيار .

مثال (5-49)

أوجد بكم طريقة يمكن اختيار فريق رياضي يتكون من ثلاث رجال وسيدتين من بين سبعة رجال وخمس سيدات .

الحل :

أولاً - يمكن اختيار السيدتين من بين الخمس سيدات بطرق عددها $\binom{5}{2}$ ، كما ويمكن اختيار الثلاث رجال من بين السبعة رجال بطرق عددها $\binom{7}{3}$ ، وبذلك طيّمكنا اختيار الفريق المطلوب بطرق عددها :

$$\binom{7}{3} \times \binom{5}{2} = \frac{7.6.5}{3.2.1} \cdot \frac{5.4}{2.1} = 350$$

وهكذا نجد أنه يوجد 350 طريقة يمكن فيها اختيار الفريق الرياضي المطلوب .

$\binom{n}{r}$ أشرنا في السابق إلى معاملات ذات الحدين وأوضحنا أن الرمز $\binom{n}{r}$ لعدنان صحيحان موجبان n, r بحيث يكون $r \leq n$ كما يلي :

$$\binom{n}{r} = \frac{n(n-1)(n-2)\dots\dots\dots(n-r+1)}{1.2.3.4\dots\dots\dots(r-1)r}$$

وتسمى هذه الأعداد بمعاملات ذات الحدين . وبالعودة إلى هذا الموضع لتوضيح نظرية ذات الحدين نجد أن هذه النظرية تعطي بالاستنتاج الرياضي للصورة العامة لمفكوك الحد $(a + b)^n$. أن نظرية ذات الحدين (Binomial Theorem) تعطي رياضياً كما يلي :

$$(a+b)^n = \sum_{r=0}^n \binom{n}{r} a^{n-r} b^r$$

$$= a^n + n.a^{n-1}b + \frac{n(n-1)}{1.2} a^{n-2}b^2 + \dots\dots\dots + nab^{n-1} + b^n$$

ويجب ملاحظة الخواص التالية لمفكوك $(a + b)^n$:

- 1- يتناقص أس a حداً بعد حد من n إلى الصفر ، ويزيد أس b حداً بعد حد من الصفر إلى n .
- 2- معامل كل حد هو $\binom{n}{k}$ حيث k هو أس أي من a أو b .
- 3- تتساوى معاملات الحدود التي تبعد عن بداية المفكوك ونهاية المفكوك بنفس المقدار .
- 4- مجموع أس a و b في كل حد هو n .
- 5- يوجد في المفكوك حدوداً عددها $n + 1$.

مثال (50-5)

أوجد مفكوك الحد الرياضي باستخدام نظرية ذات الحدين $(x+y)^5$.

الحل :

نستخدم التعريف الرياضي لنظرية ذات الحدين لإيجاد مفكوك هذا

الحد حيث :

$$\begin{aligned}(x+y)^5 &= (x)^5 + \frac{5}{1}(x)^4(y)^1 + \frac{5.4}{2!}(x)^3(y)^2 + \frac{5.4}{2!}(x)^2(y)^3 + \frac{5}{1}(x)(y)^4 + (y)^5 \\ &= x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5\end{aligned}$$

مثال (51-5)

أوجد مفكوك المقدار $(x^2 - 2y)^6$ وبعد ذلك اختصر المقدار الى

أبسط صورة .

الحل :

باستخدام نظرية ذات الحدين حيث :

$$\begin{aligned}(x^2 - 2y)^6 &= (x^2)^6 + \frac{6}{1}(x^2)^5(-2y) + \frac{6.5}{2!}(x^2)^4(-2y)^2 + \frac{6.5.4}{3.2!}(x^2)^3(-2y)^3 \\ &+ \frac{6.5}{2!}(x^2)^2(-2y)^4 + \frac{6}{1}(x^2)(-2y)^5 + (-2y)^6\end{aligned}$$

$$= x^{12} - 12x^{10}y + 60x^8y^2 - 160x^6y^3 + 240x^4y^4 - 192x^2y^5 + 64y^6$$

وهكذا باستخدام نظرية ذات الحدين نجد أنه يمكننا إيجاد مفكوك المقادير

كثيرة الحدود ، وأن هذه النظرية سوف تساعدنا في دراسة التوزيعات

الاحتمالية في الباب اللاحق .

20.5 تمارين

س1: أكتب فضاء العينة لتجربة إلقاء حجر نرد زهر ، الأول أخضر والثاني أحمر .

س2: أكتب فضاء العينة في تجربة إلقاء قطعة نقود فإذا كان الناتج صورة وأعدنا إلقاءها ثانية ، أما إذا كتابة فإننا نلقي حجر النرد الزهر .

س3: أربعة طلاب انتخبوا عشوائياً من فصل دراسي ما ، فإذا رمز بالرمز M للطالب ورمز للطالبة بالرمز F ، أكتب فضاء العينة S_1 . ثم أكتب فضاء عينة ثاني S_2 عناصره تمثل عدد الإناث المنتخبة .

س4: افرض أن A , B , C أحداثاً معينة ، والمطلوب التعبير عن وتكوين شكل فن للحاات التالية :

- (a) وقوع حدث واحد بالضبط من هذه الاحداث .
- (b) عدم وقوع أي حدث منهما .
- (c) وقوع الحدث B والحدث C وعدم وقوع الحدث A .
- (d) وقوع حدثين على الأقل منهما .

س5: بالعودة إلى السؤال الاول أوجد :

(a) أكتب عناصر الحدث A الذي يقع إذا كان المجموع أقل من 5 ثم أحسب احتمال A .

(b) أحسب احتمال الحدث B الذي يقع إذا ظهر العدد 2 على الحجر الأخضر .

(c) أحسب احتمال الحادث C الذي يقع إذا ظهر العدد 6 على أحد حجري النرد .

(d) أحسب احتمال الحادث B/C .

(e) أحسب احتمال الحادث $AUBUC$.

س5: تقدم رجلان وامرأتان بطلباتهم لمل* وظيفتين شاغرتين مختلفتين في إحدى المؤسسات . أختار المدير عشوائياً شخصاً للوظيفة الأولى وآخر للوظيفة الثانية أوجد ما يلي :

(a) فضاء العينة S .

(b) احتمال الحادث A الذي يقع إذا كان رجلاً قد شغل الوظيفة الأولى .

(c) احتمال الحادث B الذي يقع إذا كانت وظيفة واحدة فقط قد شغلت من قبل رجل .

(d) احتمال الحادث C الذي يقع إذا كانت الوظيفتان قد شغلنا من قبل امرأتين .

س6: في لعبة حظ هناك صندوق يحتوي على 500 ظرف ، 50 منها تحتوي على 50 دينار في كل ظرف ، 100 ظرف منها تحتوي على 25 دينار في كل ظرف ، 350 ظرف منها تحتوي على 10 دنانير في كل ظرف . إذا تم دفع 25 دينار يمكن سحب أحد الظروف .

(a) أحسب احتمال أن يحتوي الظرف المسحوب على 50 دينار .

(b) أحسب احتمال أن يحتوي الظرف المسحوب على أقل من 50 دينار .

س7: سحبت ورقة من شدة كوتشينة عادية عددها 52 . إذا كان A يمثل الحدث الذي يقع إذا كانت الورقة المسحوبة حمراء . B يمثل الحدث الذي يقع إذا كانت الورقة المسحوبة تحمل رقماً أكبر من 2 وأقل من 9 ، أحسب ما يلي :

(a) احتمال الحدث A .

(b) احتمال الحدث B .

(c) احتمال وقوع الحدث A بشرط إذا كان الحدث B قد وقع أي $P(A/B)$.

(d) $P(A \cup B)$.

س8: سحب كرة عشوائياً من كيس يحتوي على ثلاثة كرات حمراء (R) و 4 كرات خضراء (G) و 5 كرات صفراء (Y) أوجد :

(a) احتمال أن لا تكون الكرة المسحوبة صفراء اللون .

(b) احتمال كون الكرة المسحوبة صفراء أو حمراء .

س9: في استفتاء تلفزيوني تلقت إدارة التلفزيون 2000 جواب موزعة كما هو مبينة في الجدول (5-8) .

جدول (5-8)

المرسل	نعم	كلا	المجموع
ذكر	240	560	800
أنثى	270	930	1200
المجموع	510	1490	2000

(a) سحب ظرف عشوائياً فكانت النتيجة نعم ، ما احتمال أن يكون المرسل ذكراً .

(b) سحب ظرف إجابة عشوائياً بعد إعادة الأول فكانت الرسالة أنثى ، ما احتمال أن يكون الجواب كلا .

س10: مجموعة من الطلاب العرب 10 أردنيين و30 لبناني و10 عراقيين تقدموا لامتحان نهاية العام فنجح 3 أردنيون و 10 لبنانيون و5 عراقيون . اختير طالب عشوائياً من هذه المجموعة فكان من الناجحين . ما احتمال أن لا يكون هذا الطالب عراقياً .

س11: صندوق يحتوي على 8 كرات حمراء (R) و3 كرات بيضاء (W) و9 كرات زرقاء (B) . سحب ثلاث كرات عشوائياً ، أوجد الاحتمالات التالية :

- (a) احتمال أن تكون الكرات الثلاثة حمراء .
- (b) احتمال أن تكون كرة حمراء وكرتين بيضاء .
- (c) احتمال أن تكون كرة من كل لون .
- (d) احتمال أن تكون الأولى حمراء والثانية بيضاء والثالثة زرقاء .

س12: مجموعة عشوائية مؤلفة من n شخص ، ما احتمال أن يكون بينها أي شخصين لهما نفس تاريخ الميلاد . وبفرض أن 23 يساوي n ما قيمة هذا الاحتمال ، وما احتمال أن يكون هناك شخصان على الأقل لهما نفس تاريخ الميلاد مع العلم أن $n = 23$ أيضاً .

س13: أوجد الاحتمالات لكل من الحوادث التالية :

- (a) احتمال الحصول على عدد فردي عند رمي حجر النرد لمرة واحدة .
- (b) احتمال الحصول على مجموع 9 عند رمي حجري نرد .
- (c) احتمال أن يكون مجموع ناتج الوجهين العلويين للحجر أكبر من 10 أو أقل من 5 .

س14: فصل به 12 تلميذاً خمسة منهم ذكور والبقية إناث . فإذا اخترنا تلميذاً واحداً بطريقة عشوائية فما هو احتمال :

- (a) أن يكون التلميذ المختار فتاة .
- (b) إذا اخترنا تلميذين عشوائياً فما هو احتمال أن تكونا فتاتين .

س15: سحب كارتان من ورق اللعب فأوجد احتمال أن يكون كل منهما آس حسب الشروط التالية :

- (a) في حالة إرجاع الكارت الأول وخط السورق جيداً قبل سحب الكارت الثاني .
- (b) في حالة عدم إرجاع الكارت الأول .

س16: من مجموعة من الكرات مرقمة من 1 إلى 17 سحب كرة عشوائياً فما هو احتمال :

- (a) أن يكون الرقم المدون عليها يقبل القسمة على 2 أو 7 .
- (b) أن يكون الرقم المدون عليها يقبل القسمة على 3 أو 5 .

س17: إذا سحب كرتان عشوائياً من كيس به خمس كرات بيضاء (W) وثمانية كرات سوداء (B) فما هو احتمال :

- (a) أن تكون الكرتان المسحوبتان من لونين مختلفين .
- (b) أن تكون الكرتان المسحوبتان من لون واحد .
- (c) أن تكون واحدة على الأقل من الكرتين سوداء .

س18: يحتوي صندوق على أربع كرات بيضاء (W) وثلاث كرات سوداء (B) ويحتوي صندوق آخر على ثلاث كرات بيضاء وخمس كرات سوداء وسحب كرة من كل صندوق ، أوجد احتمال أن تكون :

(a) كل منهما بيضاء .

(b) كل منهما سوداء .

س19: إذا اشترى شخص ورقة يانصيب حيث الجائزة الأولى 50 ديناراً والجائزة الثانية 20 ديناراً باحتمال 0.001, 0.003 على الترتيب . فما هو السعر العادل الذي يدفعه لهذه الورقة .

س20: لوحظ أن متوسط المسامير الناتفة نسبة لمواصفات معينة التي تنتجها آلة محددة في مصنع هو 20% . فإذا اخترنا 10 مسامير عشوائياً من الإنتاج اليومي لهذه الآلة . أوجد احتمال وجود ما يلي :

(a) مسارين تالفين فقط .

(b) مسارين تالفين أو أكثر .

(c) أكثر من خمسة مسامير تالفة .

س21: إذا كان احتمال أن يعيش سعيد 20 عاماً آخرأ هو 60 % واحتمال أن تعيش زوجة سعيد 20 عاماً آخرأ أيضاً هو 80 % ، فما هو احتمال أن يظل الاثنان على قيد الحياة 20 عاماً .

س22: في مدينة ما إذا علم أن 40% من المواطنين لهم شعر بني اللون و 20% لهم عيون بنية اللون و 10% لهم شعر بني عيون بنية . فإذا أختير مواطن بطريقة عشوائية من هذه المدينة فأوجد كل من الاحتمالات التالية :

(a) إذا كانت عينيه بنية فما هو احتمال أن يكون شعره ليس بنياً .

(b) ما هو احتمال ان لا يكون شعره بنياً وأن لا تكون عينيه بنية .

(c) إذا كان شعره بني فما هو احتمال ان تكون ايضاً عيناه بنيتان .

س23: في فصل دراسي 12 طالب و 4 طالبات ، إذا أختير 3 طلبة من الفصل بطريقة عشوائية فما هو احتمال أن يكونوا جميعاً طلاباً .

س24: في إحدى كليات الهندسة وجد أن 35% من الطلبة رسبوا في امتحان خواص المواد ، ورسب 20% من الطلبة في امتحان الميكانيكا النظرية ، ورسب 15% في امتحان خواص المواد والميكانيكا النظرية . أختير أحد الطلبة بطريقة عشوائية أوجد ما يلي :

(a) إذا كان راسباً في خواص المواد فما هو احتمال أن يكون راسباً في الميكانيكا النظرية .

(b) ما هو احتمال أن يكون راسباً في خواص المواد والميكانيكا النظرية .

(c) إذا كان راسباً في الميكانيكا النظرية فما هو احتمال أن يكون راسباً في خواص المواد .

س25: تنتج ثلاث ماكينات A , B , C على التوالي 20% و 35% و 5% من الانتاج الكلي لمصنع انتاج الرقائق الالكترونية الخاصة بتصنيع الحاسوب ، فإذا كانت نسبة الانتاج للرقائق المعيبة لهذه الماكينات هي على التوالي 4% ، 8% ، 3% . إذا أختيرت رقاقة الكترونية بطريقة عشوائية ووجدت أنها معيبة ، أوجد احتمال أن تكون الوحدة من انتاج الماكينة B .

س26: قام رجل بزيارة زوجين لهما طفلان ، فدخل أحد الطفلين وهو ولد إلى حجرة الجلوس ، أوجد احتمال أن يكون الطفل الآخر ولداً أيضاً إذا كان :

(a) من المعلوم أن الطفل الآخر هو الأصغر .

(b) إذا لم يكن لدينا أية معلومات عن الطفل الآخر .

س27: أوجد عدد الطرق التي يمكن أن تصف بها عشرة كتب من الحجم الكبير وستة كتب من الحجم المتوسط وأربعة كتب من الحجم الصغير على أحد الرفوف بشرط أن تكون جميع الكتب ذات الحجم الواحد مرتبة معاً .

س28: أوجد كم عدد الطرق التي يمكن أن يجلس بها خمسة أشخاص في سيارة إذا كان اثنان منهم يجيدون القيادة .

س29: رجل له ستة أصدقاء وأراد أن يدعوهم الى حفلة عيد ميلاده أوجد بكم طريقة يمكن أن :

(a) يدعو أربع أصدقاء الى هذه الحفلة .

(b) يدعو خمسة أصدقاء منهم إذا كان اثنان منهم متزوجون ولا بد من حضورهما .

س30: إذا كان مطلوب من أحد المهندسين المتقدمين لشغل وظيفة ما في أحد المصانع أن يجيب على عشرة أسئلة من بين أربعة عشر سؤالاً أوجد :

(1) كم عدد الطرق التي يمكن لهذا المهندس اختيار اسئلته .

(2) بكم طريقة يمكن للمهندس اختيار الاسئلة إذا كان لا بد أن يجيب على السؤالين الاولين .

(3) بكم طريقة يمكنه الاختيار إذا الزم بالاجابة على ثلاثة أسئلة من بين الاسئلة الستة الاولى على الاقل .

س31: أحسب كلا من المقادير التالية :

$$1) \binom{19}{14} , \quad 2) \binom{13}{2} , \quad 3) \binom{30}{10}$$

س32: أختص كلاً من المقادير التالية :

1) $\frac{(n-1)!}{(n+2)!}$

2) $\frac{(n-r+1)!}{(n-r-1)!}$

3) $\frac{n!}{(n-2)!}$

4) $\frac{(n+2)!}{n!}$

س33: أوجد مفكوك كل من الحدود الآتية ثم أختصر كلاً منها :

1) $(2x + y^2)^5$

2) $(x^2 - 3y)^7$

3) $(2x^2 - y)^4$

س34: أوجد الحد الذي يحتوي على x^6 في مفكوك المقدار :

$$(2x^2 + 1/2y^3)^8$$

س35: تلقى قطعة نفود حتى تظهر الصورة لأول مرة أو تظهر الكتابة خمس مرات متتالية ، أوجد القيمة المتوقعة E لعدد مرات ألقاء القطعة .

س36: صندوق يحتوي على 10 وحدات انتاجية لاحدى المصانع من بينها ثلاث وحدات معيبة ، أختار مهندس ثلاث وحدات من الصندوق ، أوجد توقع عدد الوحدات المعيبة التي أختارها هذا المهندس .

الباب السادس

المتغيرات العشوائية والتوزيعات الاحتمالية

(Rondom Variables & Probabilities Distributions)

- 1.6 مقدمة .
- 2.6 المتغيرات العشوائية .
- 3.6 دالة التوزيع الاحتمالي .
- 4.6 التوقع الرياضي والتباين للمتغير العشوائي .
- 5.6 الإحصائيات والمعالم .
- 6.6 المتغير العشوائي المعياري .
- 7.6 التوزيعات الاحتمالية .
- 1.7.6 التوزيع ذو الحدين .
- 2.7.6 توزيع بواسون .
- 3.7.6 التوزيع الطبيعي .
- 4.7.6 منحنى التوزيع الطبيعي .
- 8.6 خواص المنحنى الطبيعي .
- 9.6 المتغير الطبيعي المعياري .
- 10.6 تحويل المتغير الطبيعي إلى متغير طبيعي معياري .
- 11.6 جداول التوزيع المعتدل المعياري .
- 12.6 تقريب توزيع ذو الحدين باستخدام التوزيع الطبيعي .
- 13.6 توزيع مربع Chi .
- 14.6 توزيع (t) .
- 15.6 توزيع (F) .
- 16.6 نمازين .

في الباب السابق تم عرض ودراسة مبادئ نظرية الاحتمالات ، ومن ضمنها تم التعرض لمفهوم فضاء العينة S وكيفية إيجاد عناصره ومجموعاته الجزئية ، وكذلك تم التطرق إلى دراسة الأحداث وكيفية حساب احتمالات وقوعها . بالإضافة إلى أنه تم بشكل سريع التعرض لمفهوم المتغير العشوائي وتوقعه الرياضي .

وسوف نقوم في هذا الباب بدراسة المتغيرات العشوائية (Random Variable) ، التي تعتبر إحدى المفاهيم الأساسية في نظرية الاحتمالات ، والاحتمالات الخاصة بها وكيفية إيجاد دوال التوزيعات الاحتمالية (Probabilities Distributions) المناظرة لها بشكل مفصل ، والتي تساعد في الحصول على النتائج التي تستخدم في الإحصاء الاستنتاجي . والذي بواسطته تتخذ القرارات الإحصائية على أسس علمية سليمة ، لذا فإن دراسة التوزيعات الاحتمالية تعتبر ذو أهمية بالغة في دراسة العديد من الظواهر والتطبيقات في الحياة العملية ، ومن أمثلة هذه التوزيعات توزيع بواسون وتوزيع ذات الحدين والتوزيع الطبيعي وغيرها من التوزيعات الهامة التي سنقوم بدراستها لاحقاً .

2.6 المتغيرات العشوائية (Random Variables)

في الباب السابق تم تعريف وتوضيح مفهوم المتغير العشوائي (Random Variable) بشكل سريع . أما في هذا الباب فسوف نقوم بدراسة المتغيرات العشوائية بالتفصيل نظراً لارتباطها الهام بالتوزيعات الاحتمالية المختلفة . نفرض أن S فضاء عينة لتجربة ما ، وأردنا تخصيص عدد معين لكل ناتج ، مثلاً طول عمر مصباح كهربائي بالأيام ، مجموع

العديدين عند إلقاء حجر الزهر وغيرها ، ويسمى مثل هذا التخصيص بشكل عام بالمتغير العشوائي ويرمز له بالرمز X ، ويعرف على أنه دالة حقيقية من S إلى مجموعة الأعداد الحقيقية R ، بحيث تكون الصورة العاكسة لأي فترة من مجموعة الأعداد الحقيقية R حدثاً في فضاء العينة S ، أي بمعنى آخر تكون قيم المتغير العشوائي مجموعة جزئية من مجموعة الإعدادات الحقيقية .

ويجب الإشارة هنا إلى أنه إذا كان فضاء العينة S فضاءً متقطعاً حيث تعرف كل مجموعة جزئية حدثاً فإن كل دالة حقيقية على S هي متغير عشوائي . ومن ناحية أخرى يمكن إثبات أنه إذا كان S فضاء غير قابل للعد فإن بعض الدوال الحقيقية ليست بمتغيرات عشوائية .

ويوجد نوعان من المتغيرات العشوائية :

(1) المتغيرات العشوائية المنفصلة (Discrete Random Variables)

من الأمثلة على هذه المتغيرات العشوائية المنفصلة عدد مرات ظهور الصورة عند رمي قطعة نقود n من المرات ، عدد حوادث المرور التي تحدث في إحدى التقاطعات خلال شهر ، عدد الزبائن الذين يصلون من شباك أحد المصارف خلال ساعة من الزمن وغيرها . ويرمز عادةً إلى المتغيرات العشوائية كما أشرنا سابقاً بالرمز X ، ويرمز للقيم التي يأخذها المتغير العشوائي بالرموز :

$$\{x_1, x_2, x_3, \dots, x_n\}$$

حيث :

$$x_1 \leq x_2 \leq x_3 \leq \dots \leq x_n$$

والاحتمالات :

$$P(X = X_1) , \quad P(X = X_2) \dots\dots\dots P(X = X_n)$$

وهكذا نجد بأن المتغيرات التي يكون نطاقها المصاحب منتهياً هي متغيرات عشوائية منفصلة أو (منقطعة) .

(2) المتغيرات العشوائية المتصلة (Continuous Random Variables)

وهو المتغير العشوائي الذي يكون مجاله المقابل غير قابل للعدد ، أي أن قيم المتغير تكون جميع القيم لفترة مثل $[a, b]$. ومن الأمثلة على هذا المتغير العشوائي المتصل أو المستمر الدخل الشهري مثلاً لمجموعة من الأسر أو درجات الحرارة لفترة زمنية معينة أو أوزان وأطوال مجموعة من الرياضيين .

فمثلاً إذا قمنا بدراسة ظاهرة الطول عند مجموعة من الرياضيين فإن المتغير العشوائي الذي يمثل طول احد الرياضيين يختلف عن طول أي رياضي آخر ، ويمكن أن يأخذ أي قيمة داخل الفترة $[a, b]$. فإذا كان طول رياضي ما هو X_1 وطول رياضي آخر هو X_2 فإنه يمكن أن نجد رياضياً ثالثاً طوله X_3 يقع بين X_1, X_2 مهما كانت القيمتان لـ X_1, X_2 قريبتين من بعضهما .

3.6 دالة التوزيع الاحتمالي

(Probability Distribution Function)

تعرف دالة التوزيع الاحتمالي للمتغير العشوائي X بأنها احتمال أن يأخذ المتغير X أقل أو يساوي قيمة معينة X ويرمز لها عادة بالرمز $f(x_i)$ ، حيث :

$$f(x) = P(X \leq x_i)$$

أي أن $f(x)$ يساوي احتمال أن يأخذ المتغير العشوائي X قيم أقل أو تساوي x_i وهي دالة في x . وهذا ما يقودنا إلى التعريف الرياضي الآتي لدالة التوزيع الاحتمالي .

فلو كان X هو متغير عشوائي معرف على فضاء العينة S بحيث تكون صورته مجموعة منتهية $X(S)$:

$$X(S) = \{x_1, x_2, x_3, \dots, x_n\}$$

ووجدنا من $X(S)$ فضاء احتمال بتعريف احتمال x_i على أنه $P(X = x_i)$ ويكتب عادة على شكل دالة $f(x_i)$ ، وتسمى هذه الدالة f المعرفة على $X(S)$ بدالة التوزيع $f(x_i) = P(X = x_i)$ (Distribution Function) ، أو دالة الاحتمال المتغير X وتعطى عادة على صورة جدول وتحقق دالة التوزيع f الشروط التالية :

$$(1) \quad \sum_{i=1}^n f(x_i) = 1$$

$$(2) \quad f(x_i) \geq 0$$

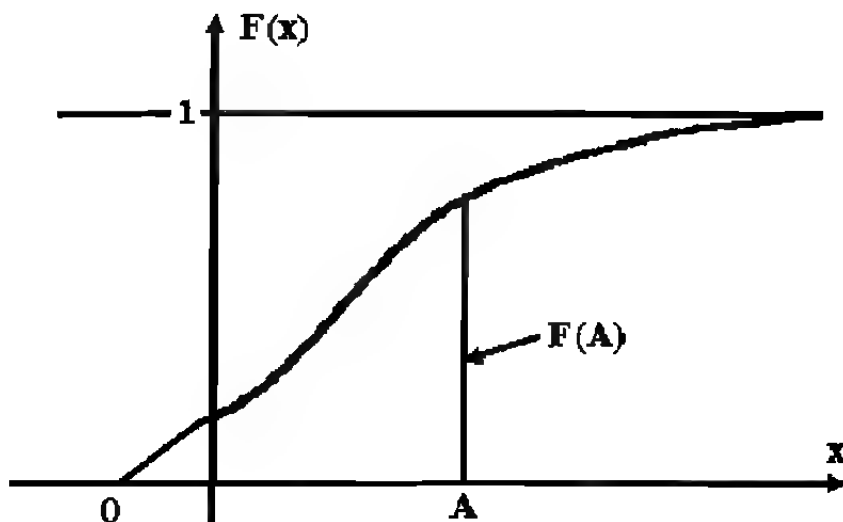
كما تجدر الإشارة إلى أنه إذا كانت $x_1 < x_2$ فإن :

$$\begin{aligned} P(x_1 \leq x \leq x_2) &= P(x \leq x_2) - P(x \leq x_1) \\ &= f(x_2) - f(x_1) \end{aligned}$$

ومن هنا يمكن أن نستنتج ما يلي :

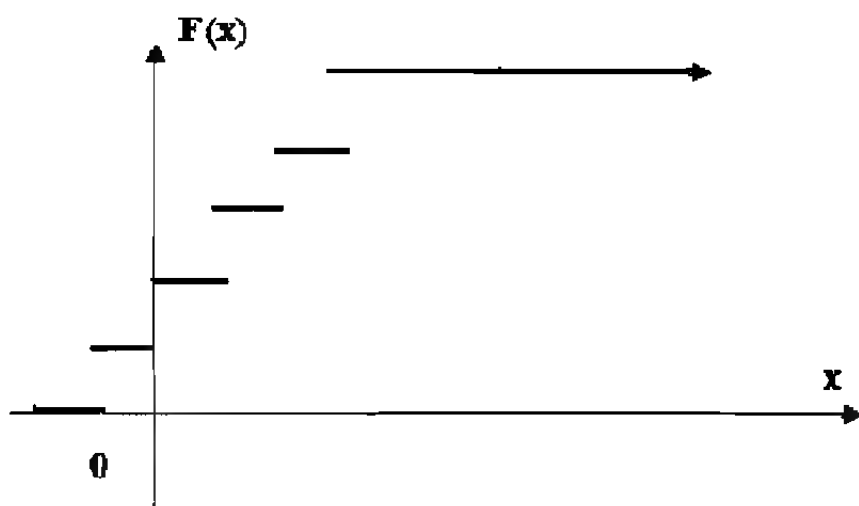
$$P(X > x) = 1 - P(X \leq x) \\ = 1 - f(x)$$

وبصفة عامة يمكن أن يأخذ منحنى دالة التوزيع الاحتمالي أحد الشكلين التاليين :



الشكل (6-1)

منحنى دالة التوزيع الاحتمالي لمتغير عشوائي متصل



الشكل (6-2)

منحنى دالة التوزيع الاحتمالي لمتغير عشوائي منفصل

نلاحظ من الشكلين أن الدالة $f(x)$ تأخذ اصغر قيمة لها وهي الصفر وتبدأ بالازدياد حتى تصل إلى أقصى قيمة لها وهي الواحد الصحيح ، ويكون المنحنى منفصلاً إذا كان المتغير العشوائي X منفصل أو متقطع ، ويكون المنحنى متصلاً إذا كان المتغير العشوائي X متصلاً أو مستمراً .

وتجدر الإشارة أنه في حالة المتغيرات العشوائية المنفصلة فإنه يمكن كتابة دالة التوزيع الاحتمالي على النحو التالي :

$$P(X \leq x) = F(x) = \sum_{x_i \leq x} f(x_i)$$

حيث أن الدالة F تعرف بدالة التوزيع المتراكمة أو دالة كثافة الاحتمال . أما في حالة المتغيرات العشوائية المتصلة فيمكن كتابة دالة التوزيع الاحتمالي على النحو التالي :

$$F(x) = \int_{-\infty}^x f(x_i) dx$$

وتسمى الدالة $f(x)$ في هذه الحالة بدالة كثافة الاحتمال ومن خواصها :

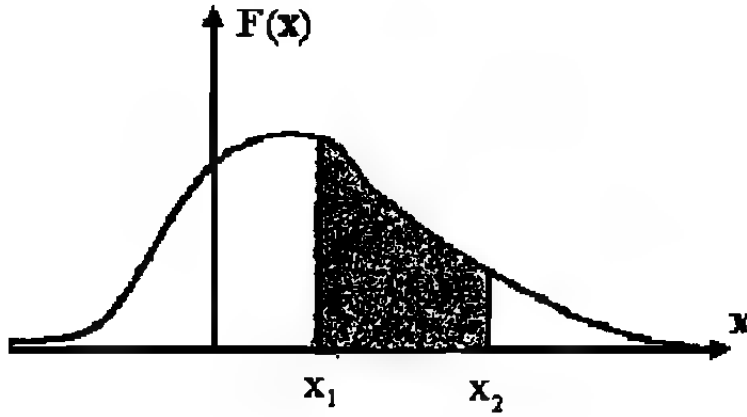
(a) أنها دالة غير سالبة أي أن $f(x)$ أكبر أو تساوي الصفر .

(b) أن تكامل لمنحنى هذه الدالة يساوي :

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

ويجب ملاحظة أن المساحة تحت المنحنى لدالة كثافة الاحتمال $f(x)$ في حالة التوزيعات الاحتمالية المتصلة تمثل احتمال أن يقع المتغير العشوائي x في فترة معينة مثل $[x_1, x_2]$ وهي المساحة المضللة والمبيّنة

في الشكل (3-6) . فإذا كانت x متغيراً عشوائياً وأخذنا المحور الأفقي ممثلاً بـ x فإنه يمكن أن نرسم منحنى يحقق المساحة الواقعة تحت المنحنى بين قيمتين معينتين مثل x_1 , x_2 للمتغير العشوائي وتساوي احتمال وقوع x بين هاتين القيمتين . ونكون بذلك قد رسمنا ما يسمى " منحنى دالة كثافة احتمال المتغير العشوائي المتصل x " ويرمز لهذه الدالة عادة بالرمز $F(x)$ وتكون المساحة الكلية تحت منحنى هذه الدالة مساوية للواحد الصحيح .



الشكل (3-6)

منحنى دالة كثافة احتمال متغير عشوائي مستمر x

مثال (1-6)

متغير عشوائي مستمر X تابع كثافته هي :

$$f(x) = \frac{2}{27}(1+x)$$

عندما تكون : $2 \leq x \leq 5$ ، والمطلوب هو :

- (1) التأكد أن $f(x)$ هو تابع كثافة احتمالي .
- (2) أوجد التوزيع الاحتمالي للمتغير العشوائي المستمر X الذي تابع كثافته $f(x)$.

(3) أوجد $P(X < 4)$.

(4) أوجد $P(3 \leq X < 4)$.

الحل:

(1) يمكن اعتبار تابع الكثافة هو :

$$f(x) = \begin{cases} \frac{2}{27}(1+x) & 2 \leq x \leq 5 \\ 0 & x < 2 \quad \text{أو} \quad x > 5 \end{cases}$$

أي أن $f(x) \geq 0$ دائماً وهذا واضح .

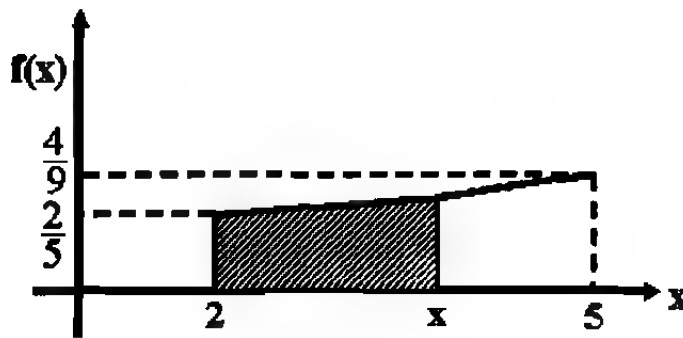
ثم أن :

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= \int_2^5 \frac{2}{27}(1+x) dx \\ &= \frac{2}{27} \left[x + \frac{x^2}{2} \right]_2^5 \\ &= \frac{2}{27} \left[\left(5 + \frac{25}{2} \right) - \left(2 + \frac{4}{2} \right) \right] = 1 \end{aligned}$$

وبالتالي فهو تابع كثافة احتمالي بناء على التعريف .

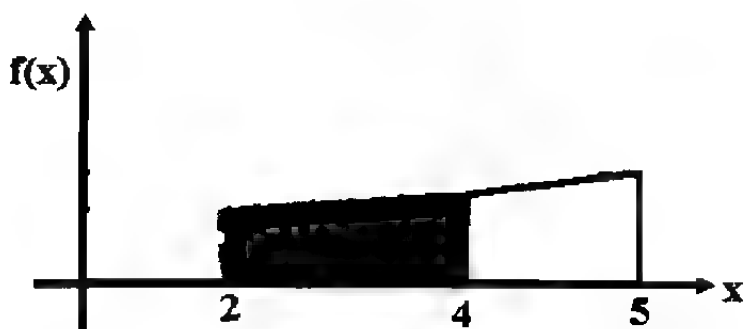
(2) تابع التوزيع الاحتمالي للمتحول العشوائي المستمر x الذي تابع كثافته $f(x)$ هو :

$$\begin{aligned}
 P(X \leq x) &= \int_{-\infty}^x f(x) dx \\
 &= \frac{2}{27} \int_2^x (1+x) dx \\
 &= \frac{2}{27} \left[x + \frac{x^2}{2} \right]_2^x = \frac{2}{27} \left[x + \frac{x^2}{2} - 4 \right]
 \end{aligned}$$



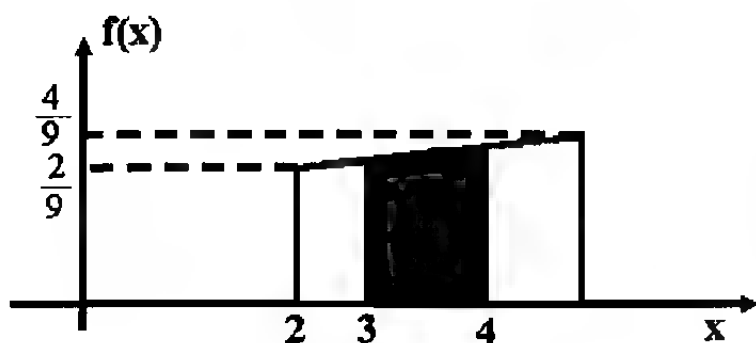
(3) إن الاحتمال المطلوب يمكن إيجاده كما يلي :

$$\begin{aligned}
 P(X \leq 4) &= \int_{-\infty}^4 f(x) dx \\
 &= \frac{2}{27} \int_2^4 (1+x) dx \\
 &= \frac{2}{27} \left[x + \frac{x^2}{2} \right]_2^4 = \frac{16}{27} = 0.59
 \end{aligned}$$



(4) إن الاحتمال المطلوب يعطى حسب التعريف :

$$\begin{aligned}
 P(3 \leq X \leq 4) &= \int_3^4 f(x) dx \\
 &= \frac{2}{27} \int_3^4 (1+x) dx \\
 &= \frac{2}{27} \left[x + \frac{x^2}{2} \right]_3^4 = \frac{9}{27} = \frac{1}{3} = 0.33
 \end{aligned}$$



إن منحني تابع كثافة في مثالنا السابق هو القطعة المستقيمة التي معادلتها :

$$f(x) = \frac{2}{27} (1+x) \quad 2 \leq x \leq 5$$

وبالتالي يمكن حساب الاحتمالات $P(x < 4)$, $P(3 \leq x < 4)$ مباشرة من حساب مساحات الأشكال الهندسية المنحرفة المظللة في الشكل . غير أن المنحنى البياني لتابع الكثافة ليس دوماً مستقيم بل هو غالباً ما يكون ذو شكل أكثر تعقيداً وهكذا فإن المساحة يجب حسابها بالتكامل المحدد .

4.6 التوقع الرياضي للمتغير العشوائي

(Expected Value of Random Variable)

في الباب السابق تم تعريف التوقع الرياضي والقيمة المتوقعة على أنه المتوسط المرجح لك قيم المتغير العشوائي ، حيث ترجح كل قيمة من المتغير باحتمالها. فإذا كان X متغيراً عشوائياً له دالة توزيع فإن التوقع أو القيمة المتوقعة للمتغير X والذي يرمز له بالرمز $E(X)$. ويعرف حسب العلاقة التالية :

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \dots + x_n f(x_n) = \sum_{i=1}^n x_i f(x_i)$$

فإذا كان X متغير عشوائي متصل فإن القيمة المتوقعة $E(X)$ له تعطى حسب العلاقة التالية :

$$E(X) = \int_{-\infty}^{\infty} X f(x) dx = \mu$$

أما إذا كان X متغير عشوائي منفصل فإن القيمة المتوقعة تعطى :

$$E(X) = \sum X_i f(x_i) = \mu$$

مثال (6-2)

متغير عشوائي X له التوزيع الاحتمالي المبين في الجدول (1-6) . أوجد التوقع الرياضي لهذا المتغير .

الجدول (1-6)

X	0	1	2	3	4
f(x)	0.1	0.3	0.3	0.2	0.1

الحل :

من التعريف نجد أن التوقع الرياضي يساوي :

$$\begin{aligned}
 E(X) &= \sum_{i=1}^n X_i f(x_i) = \mu \\
 &= 0(0.1) + 1(0.3) + 2(0.3) + 3(0.2) + 4(0.1) \\
 &= 1.9
 \end{aligned}$$

مثال (6-3)

إذا كان للمتغير العشوائي X التوزيع الاحتمالي الآتي :

$$f(x) = \left\{ \begin{array}{ll} 0.5x & , \quad 0 \leq x \leq 2 \\ 0 & , \quad \text{في فترة أخرى} \end{array} \right\}$$

أوجد التوقع الرياضي $E(X)$ لهذا التوزيع .

الحل :

أن المتغير العشوائي متصل لذا نستخدم العلاقة الآتية لإيجاد التوقع الرياضي المطلوب حيث :

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} X f(x) dx \\ &= \int_0^2 X (0.5 X) dx \\ &= 0.5 \left(\frac{x^3}{3} \right) \Big|_0^2 = 0.5 \times \frac{8}{3} \\ &= 1.34 \end{aligned}$$

بعد توضيح مفهوم التوقع الرياضي يجب علينا أن نعرف التباين والذي يرمز له بالرمز $Var(x)$ ، إذا كان X متغير عشوائي منفصل فإن التباين يعطى حسب العلاقة التالية :

$$Var(x) = \sum_{i=1}^n (x_i - \mu)^2 f(x_i)$$

أما إذا كان x متغيراً عشوائياً متصلاً فإن تباينه يعرف كما يلي :

$$Var(x) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

ويمكن استخدام العلاقة التالية في الحالتين والتي تعتبر أكثر سهولة وبساطة في حساب التباين :

$$Var(X) = E(X^2) - \mu^2$$

وقد أشرنا في الأبواب السابقة إلى أن الانحراف المعياري للمتغير العشوائي هو الجذر التربيعي الموجب للتباين أن :

$$S = \sqrt{Var(x)}$$

حيث أن :

S- هو الانحراف المعياري .

Var(x)- هو التباين .

مثال (4-6)

أحسب التوقع الرياضي والتباين والانحراف المعياري للمتغير العشوائي x الذي له التوزيع الاحتمالي المبين في الجدول (2-6) .

الجدول (2-6)

X	0	2	3	4
f(x)	0.1	0.4	0.3	0.2

الحل :

من التعريف نجد أن التوقع الرياضي يساوي :

$$E(X) = \sum_{i=1}^n X_i f(x_i) = \mu$$

$$=0(0.1)+2(0.4)+3(0.3)+4(0.2)$$

$$=2.5$$

أما التباين فيحسب حسب العلاقة التالية :

$$Var(X) = E(X^2) - \mu^2$$

ويساوي :

$$Var(x) = \sum_{i=1}^n x^2 f(x) - \mu^2$$

$$= (0)^2 \times 0.1 + (2)^2 \times 0.4 + (3)^2 \times 0.3 + (4)^2 \times 0.2 - (2.5)^2$$

$$= 7.5 - 6.25 = 1.25$$

أما الانحراف المعياري S فيساوي :

$$S = \sqrt{Var(x)}$$

$$= \sqrt{1.25}$$

$$= 1.1$$

مثال (5-6)

إذا كانت دالة كثافة الاحتمال للمتغير العشوائي X معرفة كما يلي :

$$f(x) = \left\{ \begin{array}{ll} \frac{1}{10} & , \quad 0 \leq x \leq 10 \\ 0 & , \quad \text{فترة أخرى} \end{array} \right\}$$

أوجد التوقع الرياضي والتباين والانحراف المعياري لهذا المتغير .

الحل :

نقوم بإيجاد التوقع الرياضي $E(X)$ باستخدام العلاقة التالية :

$$\begin{aligned}\mu = E(X) &= \int_{-\infty}^{\infty} X f(x) dx \\ &= \int_0^{10} \frac{1}{10} X dx \\ &= \left. \frac{X^2}{20} \right|_0^{10} = \frac{100}{20} \\ &= 5\end{aligned}$$

أما التباين فيمكن الحصول عليه عن طريق العلاقة التالية :

$$\begin{aligned}E(X)^2 &= \int_{-\infty}^{\infty} X^2 f(x) dx \\ &= \int_0^{10} \frac{1}{10} X^2 dx \\ &= \left. \frac{X^3}{30} \right|_0^{10} = \frac{1000}{3} = 33.34\end{aligned}$$

ومنه نجد التباين حيث :

$$\begin{aligned}Var(X) &= E(X)^2 - (E(x))^2 \\ &= \frac{1000}{3} - (5)^2 = \frac{25}{3} = 8.4\end{aligned}$$

أما الانحراف المعياري فيساوي :

$$\begin{aligned} S &= \sqrt{\text{Var}(x)} \\ &= \sqrt{\frac{25}{3}} \\ &= 2.9 \end{aligned}$$

أن مفهوم ومعنى التوقع الرياضي أو القيمة المتوقعة قد تم توضيحها أكثر من خلال حلول أمثلة الباب السابق .

5.6 الإحصائيات والمعالم (Statistics and Parameters)

أشرنا سابقاً أن المقياس الذي يحسب من عينة مثل الوسط الحسابي أو الانحراف المعياري وغيرها يسمى إحصائية (Statistic) . وتعتبر الإحصائية بشكل عام مقدار غير ثابت إذ الغالب أن قيمتها تتغير من عينة إلى أخرى .

أن الرقم الذي يبنى على توزيع المجتمع يسمى المعلمة (Parameter) ، فالوسط الحسابي للمجتمع هو معلمة والانحراف المعياري له معلمة أخرى وهكذا ، والمعالم هي مقادير ثابتة بالنسبة للمجتمع الواحد ولكنها تتغير من مجتمع لآخر .

وكل إحصائية نحسبها من عينة يقابلها معلمة للمجتمع وكل توزيع نظري لمتغير عشوائي يكون له معالمه الخاصة به . وقد جرت عادة الإحصائيين على استعمال الحروف الإغريقية رموزاً لمعالم المجتمع كما موضح في الجدول (3-6) حيث أن :

جدول (3-6)

الرموز المستعملة في كتابة معالم المجتمع

رمز المعلم	رمز الإحصائية (من العينة)	رمز المعلمة (للمجتمع)
الوسط الحسابي	\bar{x}	μ
الانحراف المعياري	δ	σ
راسبة أخرى	F	θ

6.6 المتغير العشوائي المعياري (Standard Random Variable)

أوضحنا سابقاً أن أي متغير يمكن كتابته باستخدام الوحدات المعيارية إذا علم كل من وسطه الحسابي وانحرافه المعياري . فإذا كانت X متغيراً عشوائياً وسطه الحسابي هو μ وانحرافه المعياري هو σ فإنه يمكننا الحصول على متغير عشوائي جديد هو \bar{x} بكتابة :

$$\bar{x} = \frac{x - \mu}{\sigma} \dots\dots\dots(1 - 6)$$

وتكون \bar{x} عندئذ هي متغير عشوائي وسطه الحسابي صفر وانحرافه المعياري هو الواحد ويطلق عليه أسم متغير عشوائي معياري (Standard Random Variable) مناظر للمتغير العشوائي x . ويجب ملاحظة أن توزيعي \bar{x} , لا يختلفان في الشكل برغم من وجود الاختلاف في نقطة الأصل وفي المقياس .

7.6 التوزيعات الاحتمالية (Probability Distribution)

نلاحظ أنه من خلال دراسة المتغيرات العشوائية (Random Variable) والتوزيعات المتعلقة بها سواء كانت متصلة أو متقطعة منفصلة أن دراسة التوزيعات الاحتمالية (Probabilities Distributions) المناظرة لها بشكل مفصل تساعد في الحصول على النتائج التي تستخدم في الإحصاء الاستنتاجي . والذي بواسطته تتخذ القرارات الإحصائية على أسس علمية سليمة ، لذا فإن دراسة التوزيعات الاحتمالية تعتبر ذو أهمية بالغة في دراسة العديد من الظواهر والتطبيقات في الحياة العملية ، ومن أهم هذه التوزيعات توزيع بواسون وتوزيع ذات الحدين والتوزيع الطبيعي وغيرها من التوزيعات الهامة التي سنقوم بدراستها بالتفصيل .

1.7.6 التوزيع ذو الحدين (Binomial Distribution)

في الحياة العملية توجد الكثير من الظواهر تكون النتائج الممكنة لها واحدة من اثنتين ، الأولى تسمى نجاحاً والثانية تسمى فشلاً . فإذا فرضنا مثلاً أن P هو احتمال النجاح فإن $q = 1 - P$ هو احتمال الفشل ، وعند تكرار التجربة n من المرات بحيث يبقى احتمال النجاح p واحتمال الفشل q ثابتين خلال مرات إعادة التجربة فإننا في كل مرة سنحصل على حالة نجاح أو على حالة فشل . أن المتغير العشوائي X الذي يمثل عدد مرات النجاح سيكون في هذه الحالة متغيراً عشوائياً منفصلاً ويأخذ القيم من $0, 1, 2, 3, \dots, n$ ويطلق عليه متغير ذو الحدين ويتوزع توزيع ذو الحدين (Binomial Distribution) بدالة كتلة الاحتمال . وهذا ما يقودنا إلى النظرية التي توضح هذا التوزيع حيث أن احتمال وقوع k من النجاحات في n من المحاولات المكررة هو :

$$b(k, n, p) = \binom{n}{k} p^k q^{n-k}$$

أن الرمز هو $\binom{n}{k}$ هو رمز التوافيق ، وقد اشرنا إليه سابقاً ويساوي :

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

حيث أن الرمز $n!$ هو مضروب العدد n ويساوي :

$$n! = n \times (n-1) \times (n-2) \times \dots \times 3 \times 2 \times 1$$

ويستخدم التوزيع ذو الحدين في كثير من الظواهر في الحياة العملية مثل تحديد الاحتمالات الخاصة بالنجاح والرسوب لعدد من الطلاب ، وإصابة أو عدم إصابة هدف معين في مباراة ما ، وعدد مواليد الذكور والإناث لمجموعة من الأسر وعدد الوحدات التي تم بيعها في عينة من المنتجات وغيرها من الظواهر .

إذا اعتبرنا أن P, n ثابت في الدالة السابقة فإن :

$$P(k) = b(k, n, p)$$

تكون توزيعاً احتمالياً منفصلاً ، حيث أن هذا التوزيع يعطى كما هو مبين في الجدول (4-6) .

الجدول (4-6)

k	0	1	2	n
P(k)	q^n	$\binom{n}{1} p q^{n-1}$	$\binom{n}{1} p q^{n-2}$	q^n

ويسمى هذا التوزيع بتوزيع نو الحدين وذلك لان الاحتمالات لقيم $n, 0, 1, 2, \dots, n$ هي الحدود المتتالية في مفكوك نو الحدين .

$$(q + p)^n = q^n + \binom{n}{1} q^{n-1} p + \binom{n}{2} q^{n-2} p^2 + \dots + p^n$$

ويطلق على هذا التوزيع في كثير من الأحيان بتوزيع العالم برنولي (Bernoulli Distribution) ، ومن خواص هذا التوزيع التوقع الرياضي والتباين والانحراف المعياري وهي مبينة في الجدول (5-6) :

جدول (5-6)

التوزيع نو الحدين	
$\mu = np$	التوقع
$\sigma^2 = npq$	التباين
$\sigma = \sqrt{npq}$	الانحراف المعياري

مثال (6-6)

ألقيت ست قطع نقود في نفس الوقت أو تم إلقاء قطعة نقود ست مرات متتالية ، أوجد ما يلي :

- احتمال وقوع صورتين بالضبط .
- احتمال وقوع أربع صور على الأقل .
- احتمال عدم وقوع صورة .
- احتمال وقوع صورة واحدة على الأقل .

الحل :

نفرض أولاً ظهور الصورة هو احتمال النجاح p ، وعليه فإن q هو احتمال الفشل حيث $p = q = 0.5$ ، و $n = 6$ وبعد ذلك نجد الآتي :

(a) أن احتمال وقوع صورتين بالضبط يعني أن $k = 2$ ، ويمكن حسابه حسب تعريف توزيع نو الحدين :

$$b(k, n, p) = \binom{n}{k} p^k q^{n-k}$$

$$\begin{aligned} b(2, 6, 0.5) &= \binom{6}{2} (0.5)^2 (0.5)^4 \\ &= 0.24 \end{aligned}$$

(b) احتمال وقوع أربع صور على الأقل يعني أن $k = 4$ أو 5 أو 6 . أي يجب علينا إيجاد ما يلي :

$$\begin{aligned} b(4, 6, 0.5) + b(5, 6, 0.5) + b(6, 6, 0.5) &= \\ &= \binom{6}{4} (0.5)^4 (0.5)^2 + \binom{6}{5} (0.5)^5 (0.5) + \binom{6}{6} (0.5)^6 \\ &= 0.34 \end{aligned}$$

(c) احتمال عدم وقوع صورة يعني الفشل في ست التجارب ويساوي :

$$q^6 = (0.5)^6 = 0.015$$

(d) احتمال وقوع صورة واحدة على الأقل يمكن حسابه من الاحتمال السابق حيث أن :

$$1 - q^6 = 1 - (0.5)^6 = 0.98$$

مثال (6-6)

أسرة بها ست أطفال ، مع فرض أن احتمال أن يكون الطفل ولد هو 0.5
أوجد الاحتمالات التالية :

- (a) أن يكون بينهم ثلاث أولاد وثلاث بنات .
(b) أن يكون عدد الأولاد أقل من عدد البنات .

الحل :

من معطيات السؤال نجد أن $n = 6$ ، $P = 0.5$ ، و $q = 1 - 0.5 = 0.5$
وبناءً على ذلك نجد المطلوب الأول :

(a) احتمال أن يكون من بين الست أطفال ثلاث أولاد وثلاثة بنات يعني :

$$P(3,6,0.5) = \binom{6}{3}(0.5)^3(0.5)^2 \\ = 0.31$$

(b) أما احتمال أن يكون عدد الأولاد أقل من عدد البنات فإن ذلك يعني :

$$P(2,6,0.5) + P(1,6,0.5) + P(0,6,0.5) \\ = \binom{6}{2}(0.5)^2(0.5)^4 + \binom{6}{1}(0.5)(0.5)^5 + (0.5)^6 \\ = 0.32$$

مثال (6-7)

أوجد التوقع الرياضي والانحراف المعياري لظهور العدد 6 لتجربة إلقاء حجر نرد الزهر 180 مرة .

الحل :

أن احتمال ظهور العدد 6 يساوي $1/6$ ، وبما أن التوزيع يتبع توزيع نو الحدين فإن التوقع الرياضي لظهور هذا العدد يعطى حسب العلاقة التالية :

$$\mu = np = 180 \times \frac{1}{6} = 30$$

أما الانحراف المعياري فيحسب كما يلي :

$$\begin{aligned}\sigma &= \sqrt{npq} \\ &= \sqrt{180 \times \frac{1}{6} \times \frac{5}{6}} = 5\end{aligned}$$

حيث أن :

$$q = 1 - p = 1 - 1/6 = 5/6$$

مثال (6-8)

إذا كان عدد المهندسين في إحدى المصانع الكبرى الخاصة بتصنيع الحاسوب يتكون من 700 مهندس و 300 مهندسة ، وأرادت إدارة المصنع اختيار وفد مكون من 4 من المهندسين والمهندسات للمساهمة في تمثيل هذا المصنع في مؤتمر علمي ما . أوجد الاحتمالات التالية :

- أن يكون بهذه اللجنة مهندسين فقط والباقي مهندسات .
- أن يكون بهذه اللجنة مهندسين على الأكثر .
- أن يكون بهذه اللجنة مهندسين على الأقل .
- متوسط عدد المهندسين بهذه اللجنة .

الحل :

أن هذا التوزيع يتبع التوزيع ذو الحدين حيث أن $n = 4$ ، وذلك بفرض أن المتغير العشوائي X يمثل عدد المهندسين الذكور بهذا الوفد ، ويأخذ القيم $0, 1, 2, 3, 4$. أن احتمال اختيار مهندس بهذا الوفد هو :

$$p = 700 / 1000 = 0.7$$

وهو يمثل احتمال نجاح تجربة ذو الحدين ، أما احتمال اختيار مهندسة بهذا الوفد فهو :

$$q = 1 - p = 300 / 1000 = 0.3$$

ونلاحظ في هذا المثال أنه من الصعب كتابة فراغ العينة S لهذه المسألة ، ولكننا باستخدام توزيع ذو الحدين نستطيع حساب الاحتمالات الخاصة بالمتغير العشوائي X الذي يتوزع توزع ذو الحدين بدالة الاحتمال :

$$f(x) = \binom{4}{x} (0.7)^x (0.3)^{4-x} , \quad x=0,1,2,3,4$$

وبناءً على ذلك نستطيع إيجاد الاحتمالات المطلوبة :

(a) أن احتمال أن يكون لهذا الوفد مهندسين فقط والباقي مهندسات ، يعني إيجاد :

$$\begin{aligned} f(2) &= \binom{4}{2} (0.7)^2 (0.3)^{4-2} \\ &= 0.27 \end{aligned}$$

(b) احتمال أن يكون بهذه اللجنة مهندسين على الأكثر يعني أنه يجب إيجاد :

$$\begin{aligned}
 P(x \leq 2) &= \sum_{x=0}^2 \binom{n}{x} P^x q^{n-x} \\
 &= \binom{4}{0} (0.7)^0 (0.3)^4 + \binom{4}{1} (0.7)^1 (0.3)^3 + \binom{4}{2} (0.7)^2 (0.3)^2 \\
 &= 0.008 + 0.076 + 0.27 \\
 &= 0.35
 \end{aligned}$$

(c) احتمال أن يكون بهذا الوفد مهندسين على الأقل يعني أنه يجب علينا إيجاد :

$$P(x \geq 2) = 1 - P(x < 2) = 1 - P(x \leq 1) = 1 - 0.083 = 0.916$$

(d) احتمال متوسط عدد المهندسين الذكور بهذا الوفد يعني أنه يجب علينا إيجاد التوقع الرياضي والانحراف المعياري للمهندسين الذكور حيث :

$$E(x) = nq = 4 \times 0.7 = 2.8$$

$$Var(x) = npq = 4 \times 0.7 \times 0.3 = 0.84$$

2.7.6 توزيع بواسون (Poisson Distribution)

كثيراً ما نصادفنا في الحياة العملية ظواهر مثل عدد من الحوادث التي تقع في تقاطع ما خلال يوم واحد ، أو عدد الأخطاء المطبعية في صفحة من صفحات كتاب ما ، أو عدد المكالمات الهاتفية في الدقيقة في أحد كبائن الاتصال أو عدد جسيمات ألفا التي يطلقها مركب نشيط إشعاعياً ، وكل هذه الظواهر تتبع توزيع احتمالي يسمى توزيع بواسون . والذي يعرف كما يلي :

$$P(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} , \quad k = 0, 1, 2, 3, \dots$$

حيث أن $\lambda > 0$ ، و e هي أساس اللوغاريتم الطبيعي وتساوي تقريباً 2.718 . ومن خواص توزيع بواسون هو ما يبينه الجدول (6-6) .

جدول (6-6)

التوزيع بواسون	
$\mu = \lambda$	التوقع
$\sigma^2 = \lambda$	التباين
$\sigma = \sqrt{\lambda}$	الانحراف المعياري

وعلى الرغم من أن توزيع بواسون له فوائد كثيرة ألا أنه يعطينا أيضاً تقريباً جيداً لتوزيع نو الحدين عندما تكون k صغيرة وبفرض p صغيرة وأن $\lambda = np$. أي بمعنى آخر عندما يكون الاحتمال P صغير جداً ويقترب من الصفر ، وعدد المحاولات n يقترب من ما لا نهاية في توزيع نو الحدين فإن هذا التوزيع يؤول إلى توزيع بواسون . ونعتبر هذه النتيجة مهمة جداً في الحياة العملية وفي كثير من التطبيقات كما سيتم توضيح ذلك من خلال حلول الأمثلة والتطبيقات القائمة .

مثال (6-9)

في إحدى المصانع وجد أن 2% من وحدات الإنتاج معيبة . أوجد احتمال أن توجد ثلاث وحدات معيبة في عينة بها 100 وحدة .

الحل :

لإيجاد الاحتمال المطلوب يمكن تطبيق توزيع نو الحدين عندما $n = 100$ و $p = 0.02$ ، وبما أن قيمة p صغيرة فإن يمكننا استخدام تقرييب بواسون عندما $\lambda = np = 2$ وعليه فإن :

$$P(3, 2) = \frac{2^3 e^{-2}}{3!} = 8(0.135) / 6 = 0.180$$

مثال (6-10)

في أحد الكتب وجد أن هناك 300 خطأ مطبعياً موزعة على صفحات كتاب به 500 صفحة . أوجد احتمال أن تحتوي صفحة معينة على :
 (a) خطأين بالضبط .
 (b) اثنان أو أكثر من الأخطاء .

الحل :

نفرض أن عدد الأخطاء في الصفحة يعبر عن عدد مرات النجاح في متتابعة توزيع ذو الحدين ، حيث أن $n = 300$ وحيث أنه يوجد 300 خطأ مطبعي واحتمال أن يظهر خطأ في الصفحة المعنية هو $P = 1/500$ وحيث أن p صغيرة فإنه يمكننا استخدام تقريب بواسون لتوزيع ذو الحدين ويكون في هذه الحالة $\lambda = np = 0.6$ وعليه نجد أن :

(a) احتمال أن يوجد خطائين بالضبط يعني أنه يجب إيجاد :

$$P(2, 0.6) = \frac{(0.6)^2 e^{-0.6}}{2!} = (0.36)(0.549) / 2 = 0.0988 = 0.1$$

(b) أن احتمال أن يوجد اثنان أو أكثر من الأخطاء يعني أن يجب إيجاد أولاً احتمال أن لا توجد أخطاء في الصفحة :

$$P(0, 0.6) = \frac{(0.6)^0 e^{-0.6}}{0!} = e^{-0.6} = 0.55$$

وثانياً أيجاد أن يوجد خطأ واحد فقط في الصفحة :

$$P(1,0.6) = \frac{(0.6) e^{-0.6}}{1!} = (0.6)(0.55) = 0.33$$

وبعد ذلك نجد احتمال اثنان أو أكثر من الأخطاء حيث أن :

$$P=1-P=1-(0.55+0.33)=0.133$$

مثال (6- 11)

وجد في إحدى المستشفيات أن نسبة الوفيات للأطفال من مرض معين هي 2% ، فإذا أخذت عينة عشوائية حجمها 300 طفل من هؤلاء الأطفال . أوجد احتمال أن يكون بهذه العينة طفل واحد متوفى على الأكثر .

الحل :

أن المتغير العشوائي X الذي يمثل عدد الوفيات يتبع التوزيع ذو الحدين ويمكننا تقريبه إلى توزيع بواسون وذلك لتوفر شروط التقريب التي أشرنا إليها سابقاً حيث :

$$\mu = np = 300 \times (0.02) = 6$$

ولإيجاد احتمال أن يكون بهذه العينة طفل واحد متوفى على الأكثر يمكن إيجاده أولاً بإيجاد :

$$P(x=0) = \frac{(0.6)^0 e^{-6}}{0!} = e^{-6} = (2.718)^{-6} = 0.0025$$

وإيجاد ثانياً أيضاً قيمة الاحتمال التالي :

$$P(x=1) = \frac{(6)^1 e^{-6}}{1!} = 6 \times e^{-6} = 6 \times (2.718)^{-6} = 0.015$$

وبناء على قيم تلك الاحتمالات يمكننا الآن إيجاد احتمال أن يكون بهذه العينة طفل واحد متوفى على الأكثر هو :

$$\begin{aligned} P(x \leq 1) &= P(x=0) + P(x=1) \\ &= 0.0025 + 0.015 \\ &= 0.018 \end{aligned}$$

3.7.6 التوزيع الطبيعي (Normal Distribution)

يوجد عدد كبير من التوزيعات النظرية التي اكتشفها الإحصائيون لغرض التحليل الإحصائي . وقد حصلوا بذلك على نتائج عظيمة في مجال الدراسات الإحصائية . ومن أهم وأشهر هذه التوزيعات المتصلة ما بالتوزيع الطبيعي (Normal Distribution) ، وقد أُشتق اسمه لأن كثيراً من الظواهر الطبيعية تأخذ شكلاً قريباً منه .

فقد لاحظ الإحصائيون أنه منذ القرن الثامن عشر أن توزيعات أخطاء المشاهدات وهي الفروق بين القيم الحقيقية والقيم المشاهدة تقترب كثيراً من هذا التوزيع . كما لاحظوا أن معظم التوزيعات البيومترية مثل ظواهر الطول

والوزن ودرجة ذكاء الإنسان وغيرها تأخذ شكلاً قريباً منه . وقد نتج عن هذا الاهتمام أن تركزت الدراسات والنظريات الإحصائية على المتغيرات التي تتوزع توزيعاً طبيعياً .

ومما زاد في فائدة دراسة التوزيع الطبيعي أنه كثيراً ما يمكن تحويل التوزيعات غير الطبيعية بطريقة أو بأخرى إلى توزيع طبيعي ، وبذلك يكون في الإمكان معالجتها باستخدام الطرق والنظريات المبينة على أساس هذا التوزيع . ودالة كثافة الاحتمال للتوزيع الطبيعي أو المنحنى الطبيعي "توزيع جاوس" تعرف كما يلي :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2(x-\mu)^2/\sigma^2}$$

حيث أن μ, σ ثابتان اختياريان وهذه الدالة تعتبر من أهم أمثلة التوزيعات الطبيعية المتصلة . ومن خواص التوزيع الطبيعي أن توقعه وتباينه وانحرافه المعياري تعطى كما هو مبين في الجدول (6-7) .

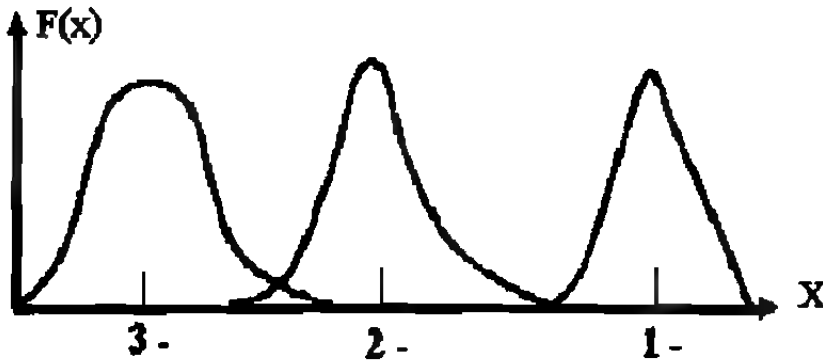
جدول (6-7)

التوزيع الطبيعي	
μ	التوقع
σ^2	التباين
σ	الانحراف المعياري

4.7.6 منحنى التوزيع الطبيعي (Normal Distribution Curve)

عندما نتكلم عن منحنى توزيع نظري فإننا نقصد منحنى دالة كثافة احتمال المتغير العشوائي لهذا التوزيع .

أن ومنحنى التوزيع الطبيعي متمائل حول خط رأسي يمر بالوسط الحسابي الذي يساوي بسبب التماثل كلاً من الوسيط والمنوال ، وللمنحنى شكل الناقوس أو الجرس ويمتد طرفاه إلى ما لا نهاية حيث يقترب طرفاه من المحور الأفقي ولكنهما لا يلتقيان معه ومع ذلك فإن المساحة تحت المنحنى تساوي الواحد الصحيح كما هو الحال في المساحة تحت منحنى دالة كثافة احتمال أي متغير عشوائي مستمر كما يبين الشكل (7-6) .

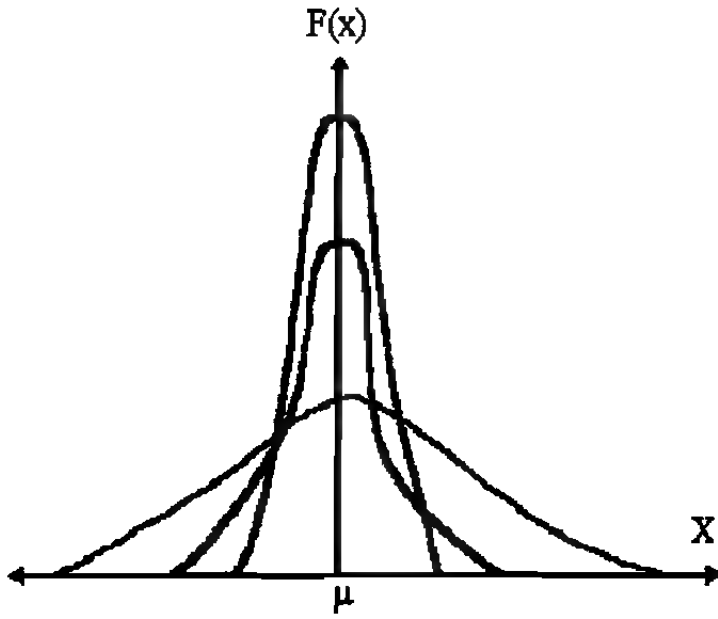


الشكل (7-6)

منحنيات معتدلة لها نفس الانحراف المعياري مع اختلافها في قيمة الوسط الحسابي

وهناك عدد لا نهائي من المنحنيات المعتدلة وهي تختلف عن بعضها البعض حسب قيمة كل من الوسط الحسابي μ والانحراف المعياري σ . وقد تتفق منحنيات معتدلة في الانحراف المعياري ولكنها تختلف في الوسط الحسابي كما في الشكل (7-6) وقد تختلف في قيمة الانحراف

المعياري وتتساوى في قيمة الوسط الحسابي كما موضح في الشكل (8-6) . مقياس الرسم مختلف في الشكلين إذ أن المساحة تحت منحنى دالة كثافة احتمال أي متغير عشوائي مستمر يجب أن يساوي الواحد .



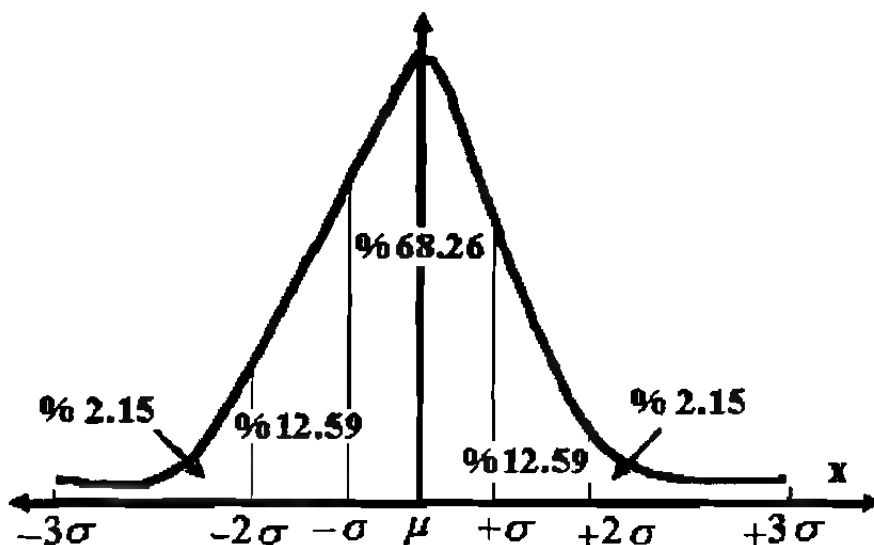
الشكل (6- 8)

منحنيات معتدلة لها نفس الوسط الحسابي مع اختلافها في قيمة الانحراف المعياري

8.6 خواص المنحنى الطبيعي (Normal Curve Properties)

مهما كانت قيمة μ و σ فإن للمنحنى الطبيعي الخواص التالية :

- (a) حوالي 68.38 % من المساحة تقع بين القيمتين $\mu - \sigma$, $\mu + \sigma$.
- (b) حوالي 95.45 % من المساحة تقع بين القيمتين $\mu - 2\sigma$, $\mu + 2\sigma$.
- (c) حوالي 99.73 % من المساحة تقع بين القيمتين $\mu - 3\sigma$, $\mu + 3\sigma$.



الشكل (6 - 9)
المساحة تحت المنحنى الطبيعي

وسنرمز للتوزيع الطبيعي الذي توقعه μ وانحرافه المعياري σ بالرمز :

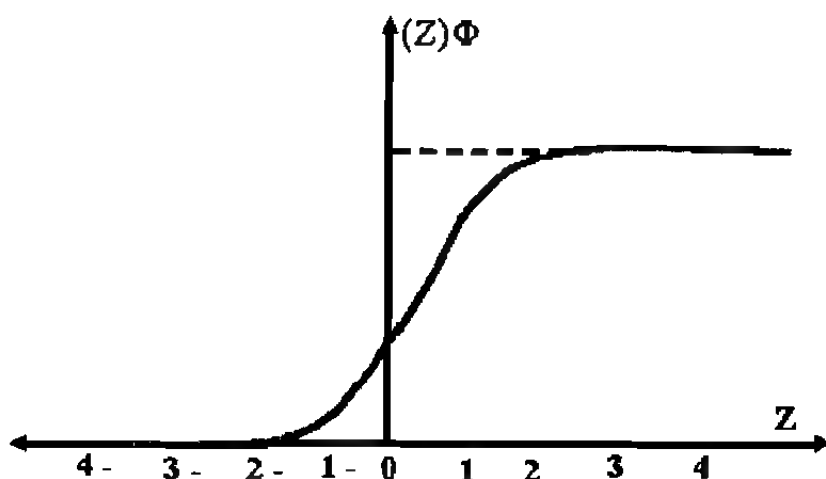
$$N(\mu, \sigma^2)$$

أي يكون له تباين هو σ^2 وبذلك يكون $N(0, 1)$ هو رمز للتوزيع المعياري المعتدل أي المتغير المعتدل الذي وسطه الحسابي يساوي صفر وانحرافه المعياري هو الواحد الصحيح . وكمثال إذا كتبنا $N(0.16, 20)$ فإننا نقصد متغيراً طبيعياً له وسط حسابي 20 وتباين 0.16 ، أي انحراف معياري يساوي 0.4 .

9.6 المتغير الطبيعي المعياري (Standard Normal Variable)

يطلق اسم المتغير المعتدل المعياري على المتغير $N(0, 1)$ ، أي على المتغير المعتدل الذي وسطه الحسابي صفر وانحرافه المعياري هو الواحد

ويرمز لهذا المتغير عادة بالرمز " Z " كما يرمز لدالة كثافة احتماله $f(Z)$ ودالة التوزيع للمتجمع بالرمز $\Phi(Z)$ كما يبين الشكل (10-6) .



الشكل (10-6)

منحنى دالة التوزيع المتجمع للمتغير المعتدل المعياري

فإذا كانت a قيمة معينة موجبة أو صفر أو سالبة فإن :

$$\phi(a) = P(X \leq a) \quad \dots\dots\dots(2-6)$$

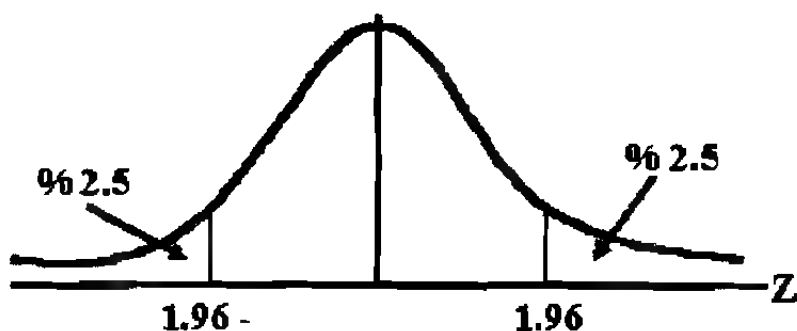
ومن الصفات العامة لتوزيع X ما يلي :

$$P(-1.96 \leq X \leq 1.96) = 0.95 \quad \dots\dots\dots(3-6)$$

$$P(-2.85 \leq X \leq 2.85) = 0.99 \quad \dots\dots\dots(4-6)$$

$$\phi(1.64) = P(X \leq 1.64) = 0.95 \quad \dots\dots\dots(5-6)$$

$$\phi(2.32) = P(X \leq 2.32) = 0.99 \quad \dots\dots\dots(6-6)$$



الشكل (11-6)

وبما أن 0.95 من المساحة تحت المنحنى $\Phi(Z)$ تقع بين القيم $X = -1.96$ و $X = 1.96$ ، فإن 0.05 من المساحة تقع خارج هاتين القيمتين . ومن تماثل التوزيع نجد أن :

القيمة 0.025 من المساحة تحت منحنى $\Phi(Z)$ يقع إلى يسار $X = 1.96$ أو أن :

$$\phi(1.96 -) = 0.025 \quad \dots\dots\dots(7 - 6)$$

والقيمة 0.025 من المساحة تقع إلى يمين $X = 1.96$ إذن 0.975 من المساحة يقع إلى يسار $X = 1.96$ أو :

$$\phi(1.96) = 0.975 \quad \dots\dots\dots(8 - 6)$$

ومن (7-6) و(8-6) نجد أن :

$$\phi(1.96-) = 1 - \phi(1.96) \quad \dots\dots\dots(9 - 6)$$

وهذه حالة خاصة لقاعدة عامة ، إذ أنه إذا كانت لـ a أي قيمة موجبة فإنه
نتيجة لتمثيل منحنى $\Phi(Z)$ حول $X = 0$ يكون :

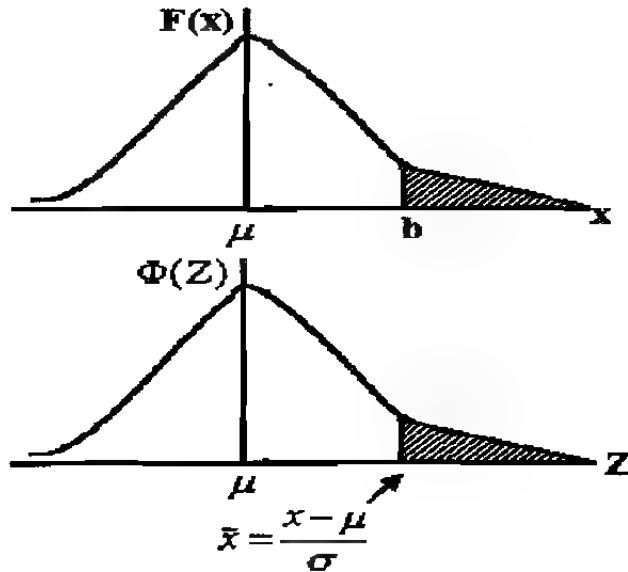
$$\phi(a-) = 1 - \phi(a) \quad \dots\dots\dots(10-6)$$

10.6 تحويل المتغير الطبيعي $N(\mu, \sigma^2)$ إلى متغير طبيعي معياري (Transform of Standard Variable to Standard Normal Variable)

إذا كانت x متغيراً طبيعياً له وسط حسابي μ وانحراف معياري σ ، فإنه يمكن تحويله إلى متغير طبيعي معياري X كما يلي :

$$\bar{x} = \frac{x - \mu}{\sigma}$$

وفي هذه الحالة فإن كثافة دالة احتمال المتغير العشوائي x أي $F(x)$ تتحول
إلى $\Phi(X)$ كما يبين الشكل (12- 6) .



الشكل (12- 6)

تحويل المتغير المعتدل إلى متغير معتدل معياري

وإذا كانت b قيمة معينة من قيم x فإن :

$$\begin{aligned} P(x \leq b) &= P(x - \mu, b - \mu) \\ &= P\left(\frac{x - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{b - \mu}{\sigma}\right) \end{aligned}$$

أو:

$$P(X \leq b) = \phi\left(\frac{b - \mu}{\sigma}\right) \dots\dots\dots(11-6)$$

وإذا كانت :

$$\begin{aligned} x > b &= 1 - P(x \leq b) \\ &= 1 - \phi\left(\frac{b - \mu}{\sigma}\right) \dots\dots\dots(12-6) \end{aligned}$$

وإذا كانت $b > a$ كما في الشكل (6-13) فإن :

$$\begin{aligned} P(a \leq x \leq b) &= \text{المساحة المظللة في الشكل} \\ &= \text{المساحة إلى يسار } b - \text{المساحة إلى يسار } a \\ &= \phi\left(\frac{b - \mu}{\sigma}\right) \cdot \phi\left(\frac{\mu - a}{\sigma}\right) \quad (14-6) \end{aligned}$$

ويبين الجدول (6-8) قيم التوزيع المتجمع الطبيعي المعياري $\Phi(Z)$ أو قيم المساحة تحت منحنى $\Phi(Z)$ من $-\infty$ إلى Z .

جدول (8-6)

قيم المساحة تحت منحنى $\Phi(Z)$ من $-\infty$ إلى Z

0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.005	Z
0.536	0.532	0.528	0.524	0.52	0.516	0.512	0.508	0.504	0.500	0.005
0.575	0.571	0.568	0.564	0.56	0.556	0.552	0.548	0.543	0.54	0.1
0.614	0.61	0.606	0.603	0.598	0.594	0.591	0.587	0.583	0.579	0.2
0.651	0.648	0.644	0.641	0.637	0.633	0.63	0.626	0.621	0.618	0.3
0.688	0.684	0.681	0.677	0.674	0.670	0.666	0.663	0.659	0.655	0.4
0.723	0.719	0.716	0.712	0.709	0.706	0.702	0.698	0.695	0.691	0.5
0.755	0.752	0.748	0.745	0.742	0.739	0.736	0.732	0.729	0.726	0.6
0.785	0.782	0.779	0.776	0.772	0.770	0.767	0.764	0.761	0.758	0.7
0.812	0.810	0.808	0.805	0.802	0.799	0.796	0.793	0.791	0.788	0.8
0.839	0.837	0.834	0.831	0.829	0.826	0.823	0.821	0.818	0.816	0.9
0.862	0.86	0.858	0.855	0.853	0.850	0.848	0.846	0.844	0.841	1.0

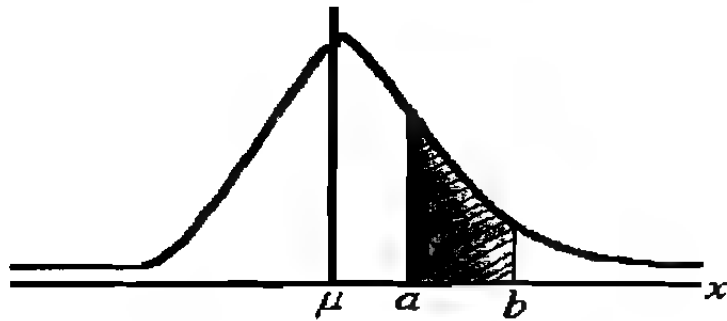
11.6 جداول التوزيع المعتدل المعياري

(The Standard Normal Variable Table)

توجد جداول مختلفة تعطي بعض أو كل قيم الدوال الآتية للمتغير الطبيعي المعياري عند قيم معينة لهذا المتغير ومن أهم خواصها :

- الإحداثي الرأسي عند النقطة Z أي $f(Z)$.
- المساحة تحت المنحنى المعتدل المعياري إلى يسار الإحداثي الرأسي عند النقطة Z ، أي $\Phi(Z)$.
- المساحة تحت المنحنى المعتدل المعياري بين الإحداثيين الرأسيين المارين بنقطة الأصل والنقطة Z (حيث X أكبر من الصفر) .

أن الجدول الخاصة بالتوزيعات المختلفة مبينة في الملحق الموجود نهاية هذا الكتاب .



الشكل (6-13)

مثال (6-12)

إذا كانت درجات الحرارة في مدينة عمان خلال شهر مارس تتبع التوزيع الطبيعي بمتوسط قدره 20°C درجة مئوية وانحراف معياري يساوي 3 درجات مئوية . أوجد احتمال أن تكون درجة الحرارة بين 12 ، 26 درجة مئوية خلال شهر مارس .

الحل :

يجب علينا أيجاد الاحتمال التالي :

$$\begin{aligned}
 P(12 < X < 26) &= P\left(\frac{12 - \mu}{\sigma}\right) < P\left(\frac{X - \mu}{\sigma}\right) < P\left(\frac{26 - \mu}{\sigma}\right) \\
 &= P\left(\frac{12 - 20}{3} < Z < \frac{26 - 20}{3}\right) \\
 &= P(-2.67 < Z < 2) \\
 &= P(Z < 2.00) - P(Z < -2.67) \\
 &= 0.977 - 0.0038 \\
 &= 0.973
 \end{aligned}$$

ويجب الإشارة هنا إلى أنه تم الحصول على هذه النتيجة باستخدام الجدول رقم (1-1) الموجود في نهاية هذا الكتاب .

مثال (6-13)

- تقيم 100 مهندس لامتحان خاص بالكفاءة في إحدى المصانع ، فوجد أن درجات المهندسين المتقدمين تتبع التوزيع الطبيعي $N(25,70)$ أوجد :
- (a) عدد المهندسين الحاصلين على درجات من 66 إلى 76 درجة .
 - (b) عدد المهندسين الحاصلين على درجات أكثر من 80 درجة .
 - (c) عدد المهندسين الحاصلين على درجات أقل من 62 درجة .

الحل :

لإيجاد المطلوب نقوم بتحويل متغير التوزيع الطبيعي إلى متغير توزيع طبيعي معياري حيث :

(a) عدد المهندسين الحاصلين على درجات من 66 إلى 76 درجة يمكن حسابه كما يلي :

$$\begin{aligned} P(66 < X < 76) &= P\left(\frac{66 - \mu}{\sigma}\right) < P\left(\frac{X - \mu}{\sigma}\right) < P\left(\frac{76 - \mu}{\sigma}\right) \\ &= P\left(\frac{66 - 70}{5} < Z < \frac{76 - 70}{5}\right) \\ &= P(-0.80 < Z < 1.20) \\ &= P(Z < 1.20) - P(Z < -0.80) \\ &= 0.2119 - 0.8849 \\ &= 0.06730 \end{aligned}$$

من الجدول رقم (1-1) حصلنا على القيمة لهذا الاحتمال ، أي باستخدام جداول التوزيع الطبيعي المعياري الموجودة في نهاية هذا الكتاب ، وبناءً عليه نجد عدد المهندسين الذين تتراوح درجاتهم من 66 إلى 76 درجة :

$$100 \times 0.673 = 68 \text{ مهندس}$$

(b) عدد المهندسين الحاصلين على درجات أكثر من 80 درجة يعني :

$$\begin{aligned} P(80 < X) &= P\left(\frac{\mu - 80}{\sigma} < \frac{X - \mu}{\sigma}\right) \\ &= P\left(Z > \frac{80 - 70}{5}\right) \\ &= P(Z > 2.00) \\ &= 1 - P(Z \leq 2.00) \\ &= 1 - 0.9773 \\ &= 0.0227 \end{aligned}$$

أذن عدد المهندسين الحاصلين على درجات أكثر من 80 درجة هو :

$$100 \times 0.0227 = 2.27 = 3 \text{ مهندس}$$

(c) عدد المهندسين الحاصلين على درجات أقل من 62 درجة هو :

$$\begin{aligned} P(62 < X) &= P\left(\frac{\mu - 62}{\sigma} < \frac{X - \mu}{\sigma}\right) \\ &= P\left(Z > \frac{62 - 70}{5}\right) \\ &= P(Z > -1.60) \\ &= 0.0548 \end{aligned}$$

أذن عدد المهندسين الحاصلين على درجات أقل من 62 درجة هو :

$$100 \times 0.0548 = 5.48 = 6 \text{ مهندس}$$

12.6 تقريب توزيع نو الحدين باستخدام التوزيع الطبيعي (Approximation of Binomial by Normal Distribution)

إذا كان X متغير عشوائي له توزيع نو حدين وكانت n كبيرة جداً واحتمال النجاح P يقترب من 0.5 فإنه يمكن تقريب دالة الاحتمال التالية :

$$f(x) = \binom{n}{x} p^x q^{n-x}$$

بدالة الاحتمال للتوزيع الطبيعي المعياري حيث أن المتوسط الحسابي والانحراف المعياري :

$$\mu = np, \quad \sigma = \sqrt{npq}$$

وبذلك يكون المتغير العشوائي $\frac{x - np}{\sqrt{npq}}$ يتوزع تقريباً بالتوزيع الطبيعي المعياري $N(0, 1)$. والمثال (14-6) يوضح هذا التقريب .

مثال (14-6)

أوجد احتمال ظهور الصورة من 40 مرة إلى 50 مرة ، إذا ألقيت قطعة نقود 100 مرة .

الحل :

التجربة هنا تتبع توزيع نو الحدين ، وبما أن قيمة n كبيرة ، واحتمال النجاح p يساوي 0.5 فإنه يمكن تقريب توزيع نو الحدين باستخدام التوزيع الطبيعي المعياري حيث :

$$\mu = np = 100 \times 0.5 = 50$$

$$\sigma \sqrt{npq} = \sqrt{100 \times 0.5 \times 0.5} = \sqrt{25} = 5$$

فإذا فرضنا أن X متغير عشوائي يمثل عدد مرات ظهور الصورة ، فإن الاحتمال المطلوب هو :

$$P = (40 < X < 50)$$

ولكن حيث أنه تم تقريب توزيع منفصل باستخدام توزيع احتمال متصل فإنه يجب احتساب معامل تصحيح وذلك بطرح وإضافة 0.5 ، أي أن الاحتمال المطلوب باستخدام تقريب التوزيع الطبيعي المعياري هو :

$$P = (39.5 < X < 50.5)$$

حيث X تتوزع تقريباً التوزيع الطبيعي $N (25 , 50)$ ، ويكون :

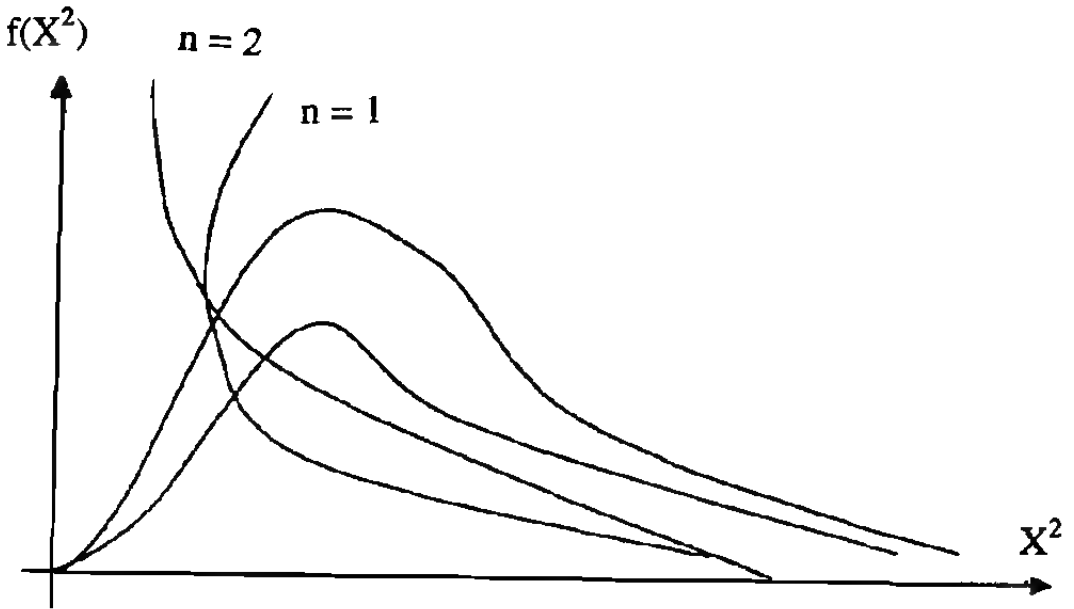
$$\begin{aligned} P &= (39.5 < X < 50.5) = P\left(\frac{39.5-50}{5} < Z < \frac{50.5-50}{5}\right) \\ &= P(-2.10 < Z < 0.10) \\ &= P(0.10 < Z) - P(-2.10 < z) \\ &= 0.0179 - 0.5398 = 0.5219 \end{aligned}$$

13.6 توزيع مربع Chi (Chi Square Distribution)

إذا أخذت عينة عشوائية مثل $X_1, X_2, X_3, X_4, \dots, X_n$ من مجتمع له توزيع طبيعي $N (\mu, \sigma^2)$ فإن المتغير العشوائي :

$$Z = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2}$$

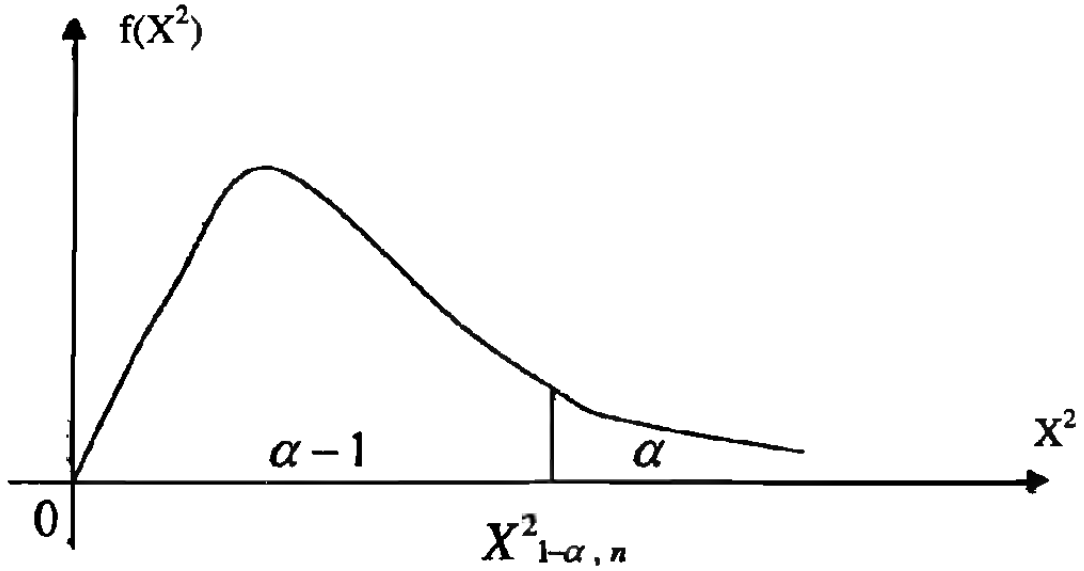
له توزيع X^2 أو ما يسمى كـ χ^2 (Chi Square) بدرجات حرية n من المتغيرات العشوائية ، ويمكن تعريفه بمعنى أبسط أنه توزيع مجموع مربعات n من المتغيرات العشوائية المستقلة المعيارية . ويطلق على n درجة الحرية لهذا التوزيع ، والتي تدل على عدد المتغيرات المستقلة الداخلة في المجموع ومن ذلك يمكن ملاحظة أن المتغير العشوائي الذي له توزيع X^2 (Chi Square) متغير غير سالب ، ومنحنى دالة كثافة احتمالية تبدأ من الصفر حيث تحدد درجات الحرية كما يبين الشكل (14-6) .



الشكل (14-6)

ويلاحظ أن المنحنيات التي درجات الحرية لها أكبر من اثنان تمس المحور الأفقي عند نقطة الأصل ثم يرتفع المنحنى حتى يصل إلى قمته العظمى ويعود إلى النزول ليمس المحور الأفقي عند ما الملا نهاية . ويعتبر توزيع X^2 (Chi Square) توزيعاً ملتوي إلى اليمين ويقترّب من التماثل كلما زادت درجات الحرية . أن أهمية توزيع X^2 (Chi Square) تكمن في استخداماته الكثيرة في الاختبارات الإحصائية وتكوين فترات الثقة الخاصة بالتباين .

إذا كانت $f(x^2)$ هي دالة التوزيع الاحتمالي للمتغير العشوائي الذي يتبع توزيع X^2 (Chi Square) ، فإن جدول توزيع مربع كاي (2-2) الموجود في نهاية هذا الكتاب يعطي قيم x^2 المقابلة لقيم معينة $f(x^2) = 1-\alpha$ ، حيث أن $1 > \alpha > 0$ تمثل احتمال معين وذلك لقيم مختارة من $1-\alpha$ ، كما يبين الشكل (15-6) .



الشكل (15-6)

ولدرجات الحرية من 1 إلى 30 ويرمز لقيمة X^2 الجدولية عند الاحتمال 1- α ودرجات الحرية n بالرمز $X^2_{1-\alpha, n}$ ، فمثلاً إذا كانت :

$$1 - \alpha = 0.70$$

ودرجات الحرية :

$$n = 17$$

فإن :

$$X^2_{0.70, 17} = 19.5$$

وكذلك :

$$X_{0.025,11}^2 = 3.82$$

أما إذا كانت درجات الحرية أكبر من 30 فإن يمكن إثبات أن المتغير العشوائي $\sqrt{2X^2} - \sqrt{2n-1}$ له توزيع يقترب من التوزيع الطبيعي المعياري .

مثال (6-15)

أوجد قيمة $X_{0.99,10}^2$.

الحل :

نجد قيمة الدالة حيث :

$$f(X^2) = 1 - \alpha = 0.99, n = 10$$

باستخدام الجدول (2-2) والنظر إلى درجات الحرية والتي تساوي 10 والمساحة 0.99 على الخط الأفقي في جدول توزيع X^2 نجد أن نقطة التقاطع هي 23.20 وهي القيمة المطلوب إيجادها أي أن :

$$X_{0.99,10}^2 = 23.20$$

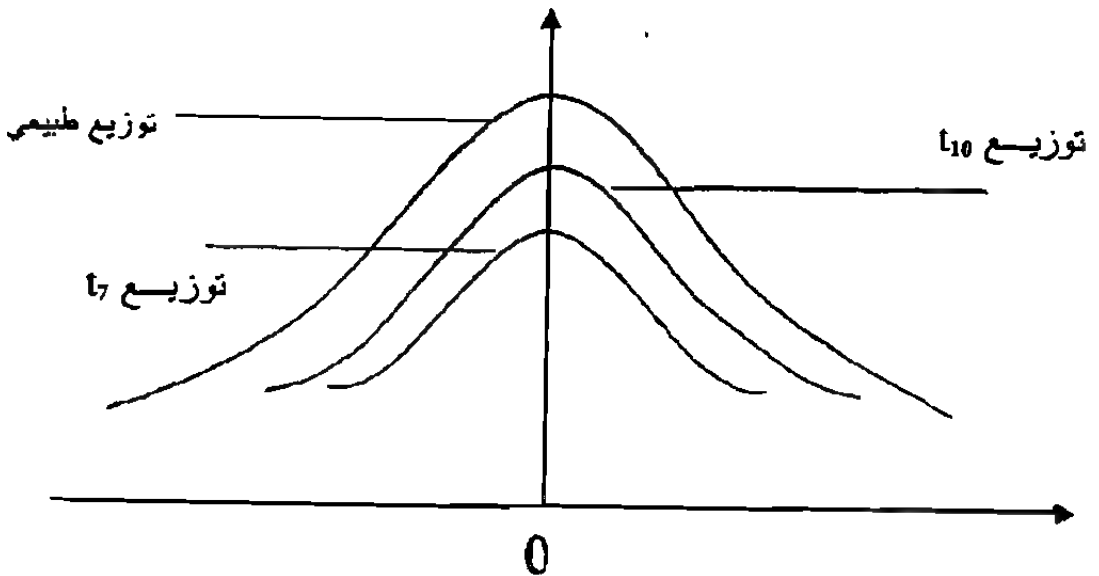
14.6 توزيع t (t - Distribution)

إذا كانت $X_1, X_2, X_3, X_4, \dots, X_n$ متغيرات عشوائية معيارية مستقلة فإن المتغير العشوائي $\sum X^2$ له توزيع X^2 (Chi Square) بـدرجات حرية n كما اشرنا في البند السابق ، وإذا كانت Z متغير عشوائي

آخر له التوزيع $N(0, 1)$ ومستقل عن المتغير $\sum X^2$ ، فإن المتغير العشوائي :

$$\frac{Z}{\sqrt{\sum_{i=1}^n X_i^2}}$$

يسمى متغير t ويتوزع توزيع t بدرجات حرية n ، ويرمز له بالرمز t_n ، أما دالة كثافة الاحتمال لهذا التوزيع فتكون متماثلة حول الصفر ولا منحني يشبه منحني دالة كثافة الاحتمال للتوزيع الطبيعي المعياري لكن الأخير يكون أعلى عند الصفر ، أي أن منحني دالة كثافة الاحتمال للتوزيع t_n أكثر تحديباً ، ويلاحظ أنه كلما زادت قيمة n كلما اقترب منحني التوزيع t_n من منحني التوزيع الطبيعي المعياري حيث يكون الشكلين متقاربين إلى حد كبير عندما تكون n أكبر أو تساوي 30 ، وفي هذه الحالة يستخدم التوزيع الطبيعي المعياري كتقريب لتوزيع t عندما تكون درجات الحرية n أكبر من 30 كما يبين الشكل (6-16) .



الشكل (6-16)

وتعطى قيمة t عند درجة الحرية n واحتمال $1-\alpha$ والتي يرمز لها بالرمز $t_{1-\alpha, n}$ من الجدول (3-3) الموجود في نهاية هذا الكتاب ، فمثلاً نجد أن
قيمة :

$$t_{0.95, 10} = 1.812 \quad , \quad t_{0.975, 14} = 2.145$$

ويلاحظ أن توزيع t متماثل حول $t = 0$ ومن التماثل نجد أنه إذا كانت t موجبة فإن :

$$P(X \geq -t) = P(X \leq t)$$

ويعني ذلك :

$$P(X \geq -t) = 1 - P(X \leq t)$$

فإذا كانت :

$$P(X \leq t) = 1 - \alpha$$

فإنه يمكن كتابة العلاقة السابقة على الصورة التالية :

$$t_{\alpha, n} = -t_{1-\alpha, n}$$

حيث أن $t_{\alpha, n}$ هي قيمة t المقابلة لاحتمال α ودرجات الحرية n ولذلك فإنه يمكن استخدام علاقة التماثل لإيجاد هذه القيمة فمثلاً :

$$t_{0.01, 10} = -t_{0.99, 10} = -2.764$$

وبالتالي فإن هذه العلاقة تفيدنا عادة في إيجاد قيم $t_{\alpha,n}$ عندما تكون قيمة α صغيرة وغير موجودة بالجدول التي أشرنا إليها سابقاً .

مثال (6-16)

أوجد قيمة $t_{0.025,21}$ ، $t_{0.99,12}$.

الحل :

نجد من الجدول (3-3) في نهاية هذا الكتاب أن قيمة :

$$t_{0.99,12} = 2.68$$

أما قيمة t في هذه الحالة فألنا نستطيع إيجادها عن طريق العلاقة السابقة حيث أن قيمة α صغيرة وغير موجودة في الجدول لذا فألنا نستخدم مكملتها :

$$t_{0.025,21} = -t_{0.975,21} = -2.080$$

15.6 توزيع F (F- Distribution)

ويعتبر هذا التوزيع من التوزيعات الاحتمالية الهامة التي تستخدم في اختبار الفرضيات ، ويمكننا تعريفه على النحو التالي :

إذا كانت X متغيراً عشوائياً له توزيع X^2 (Chi Square) بـدرجات حرية n_1 ، وكانت Z متغير عشوائي آخر مستقل عن X وله دالة توزيع احتمالي (Chi Square) أيضاً بـدرجات حرية n_2 ، فإن المتغير العشوائي :

$$\frac{Xn_1}{Zn_2}$$

له توزيع F بدرجات حرية n_1 , n_2 ، حيث يوجد الجدول (4-4) يعطي قيم F لهذا التوزيع التي تحقق احتمال أن يكون هذا المتغير العشوائي أصغر منها وذلك لقيم n_1 , n_2 المختلفة . وتكثر استخدامات هذا التوزيع في الاختبارات الإحصائية المختلفة .

مثال (6-17)

$$. F_{0.95, 12, 7} \quad , \quad F_{0.025, 20, 15} \text{ أوجد قيمة}$$

الحل :

نجد من الجدول (4-4) أن قيمة :

$$F_{0.95, 12, 7} = 3.57$$

أما قيمة :

$$F_{0.025, 20, 15} = \frac{1}{F_{0.957, 15, 20}} = \frac{1}{2.76} = 0.36$$

حيث أن القيمة السابقة تم حسابها بالعلاقة التالية :

$$F_{\alpha, n_1, n_2} = \frac{1}{F_{1-\alpha, n_1, n_2}}$$

16.6 تمرين

س1: إذا كان المتغير X عبارة عن عدد الصور التي تظهر في تجربة إلقاء قطعة من النقود مرتين متتاليتين . أوجد قيم هذا المتغير العشوائي ودالة توزيعه الاحتمالي .

س2: أوجد التوقع والتباين والانحراف المعياري للتوزيع المبين في الجدول (10-6) :

جدول (10-6)

X_i	-2	0	1	3	3
$f(X_i)$	4	1	1	4	2

س3: أوجد التوزيع والتوقع والتباين والانحراف المعياري للمتغير X في تجربة إلقاء قطعة نقود خمسة مرات ، إذا اعتبر أن X يدل على عدد مرات ظهور الصورة .

س4: صندوق يحتوي على 8 وحدات إنتاج من بينها وحتان معيبتان ، إذا اختار رجل 3 وحدات من الصندوق . أوجد توقع عدد الوحدات المعيبة التي اختارها .

س5: في تجربة اختبار نوع معين من إطارات السيارات وجد أن 15% من هذه الإطارات تفشل في اجتياز هذا الاختبار . فإذا أخذت عينة عشوائية من 20 سيارة فما هو احتمال أن يكون من 5 إلى 8 سيارات تفشل إطاراتها في اجتياز هذا الاختبار .

س6: يلقي لاعب أربعة قطع نقود حيث يكسب خمسة دنانير إذا ظهرت الصورة ثلاث مرات وثلاثة دنانير إذا ظهرت صورتان ، ودينارين إذا ظهرت صورة واحدة ، وبالمقابل فإنه يخسر خمسة عشر ديناراً إذا ظهرت الكتابة ثلاث مرات ، أوجد قيمة هذه اللعبة بالنسبة للاعب .

س7: إذا كان X متغيراً عشوائياً له التوزيع التالي :

$$f(x) = \begin{cases} \frac{1}{12} & 0 \leq x \leq x \\ 0 & \text{أخرى} \end{cases}$$

س8: إذا كان X متغير عشوائي يتوزع بدالة كثافة الاحتمال التالية :

$$f(x) = \frac{1}{8} , \quad 1 \leq X \leq 5$$

فوضح أن هذه الدالة هي دالة احتمال ثم أوجد ما يلي :

a) $P(2 < X < 4)$.

b) $E(X)$, $Var(X)$.

c) $f(X)$.

س9: إذا كانت حوادث المرور في إحدى تقاطعات طرق مدينة ما تحدث بمعدل أربعة حوادث في الأسبوع . أوجد احتمال وقوع خمس حوادث فقط في هذا التقاطع في أسبوع معين .

س10: إذا كان احتمال أن يصيب رجل هدف ما في مباراة خاصة بالرمائية هو 0.20 ، فإذا أطلق رجل سبع مرات فما هو احتمال أن يصيب الهدف على الأقل مرتين ، وكم مرة يجب أن يطلق الهدف لكي يكون احتمال أن يصيب الهدف على الأقل مرة واحدة أكبر من 0.65 .

س11: ألقي حجر نرد زهر 200 مرة أوجد احتمال أن يظهر الرقم 6 من 31 إلى 40 من المرات بما في ذلك 31 و 40 .

س12: إذا كان المتغير العشوائي يتوزع توزيع ذو الحدين حيث أن $n = 6$ و $P = 0.4$ ، أوجد ما يلي :

(a) $P(X = 2)$.

(b) $P(X \geq 1)$.

(c) $P(1 < X \leq 4)$.

س13: صندوق به ثلاث كرات حمراء وكرتان سوداء إذا سحبنا كرة ثم أعيدنا ثلاث مرات من هذا الصندوق . أوجد الاحتمالات التالية :

(a) أن تكون الكرة المسحوبة كرة حمراء .

(b) أن تكون كرتان من اللون الأحمر .

(c) على الأقل كرة حمراء .

س14: ينتج أحد المصانع أنواع معينة من الأقفال يوجد بها 6% معيبة ، أوجد توقع عدد الأقفال المعيبة وانحرافها المعياري في تشغيلة من 3800 قفل من هذا المصنع .

س15: إذا كان المتغير العشوائي X يتبع التوزيع الطبيعي $N(120, 46)$ ، أوجد الاحتمالات التالية :

(a) $P(107 < X < 111)$.

(b) $P(X > 113)$.

س16: في إحدى الامتحانات وجد أن الدرجات عبارة عن متغير طبيعي بتوقع 86 وانحراف 17 حيث يأخذ 10% من الطلبة الأوائل بالترتيب العلامة A ويأخذ 5% من الطلبة الحاصلون على أقل الدرجات بالترتيب العلامة B ، أوجد :

(1) أقل درجة لكي يحصل الطالب على العلامة A .

(2) أقل درجة يحصل عليها الطالب لكي يعتبر ناجحاً .

س17: إذا كان X متغيراً عشوائياً يتبع توزيع بواسون بمعدل أربعة ، أوجد الاحتمالات التالية :

(1) $P(X \leq 3)$.

(2) $P(X \geq 4)$.

(3) $P(2 < X < 6)$.

س18: 250 خطأ مطبعياً موزعة توزيعاً عشوائياً في كتاب به 180 صفحة ، أوجد احتمال أن تحتوي صفحة معينة على :

(a) خطئين أو أكثر .

(b) خطئين .

(c) خطأ واحداً .

(d) صفر من الأخطاء .

س19: افترض أن 4% في المتوسط من الأطفال يكتبون باليد اليسرى (العسر) ، أوجد احتمال أن يوجد 5 أطفال أو أكثر من الأطفال العسر من عينة بها 100 طفل .

س20: ظهر دواء جديد لمعالجة مرض سرطان الدم معدل نجاحه 80 %
 أعطي هذا الدواء لـ 15 مريضاً بسرطان الدم . ما احتمال شفاء 12
 منهم ، ثم ما احتمال شفاء 12 منهم على الأقل .

س21: يصيب أحد لاعبي كرة السلة 80% من رمياته من خط الرمية
 الحرة . ما احتمال أن يسجل إصابتين من أربعة رميات حرة .

س22: أفرض أن X متغير عشوائي وأن له التوزيع الطبيعي المعياري ،
 أوجد ما يلي :

(a) $P (-0.81 \leq X \leq 1.13)$.

(b) $P (0.53 \leq X \leq 2.03)$.

(c) $P (X \leq 0.73)$.

س23: أوجد قيمة كل من القيم للتوزيعات التالية :

a) $t_{0.95, 7}$ ، $t_{0.05, 19}$ ، $t_{0.975, 11}$

b) $X^2_{0.025, 14}$ ، $X^2_{0.99, 9}$ ، $X^2_{0.01, 22}$

c) $F_{0.05, 11, 7}$ ، $F_{0.975, 14, 1}$ ، $F_{0.99, 15, 8}$

جداول التوزيعات الاحتمالية المختلفة

جدول (1-1) التوزيع الطبيعي المعياري

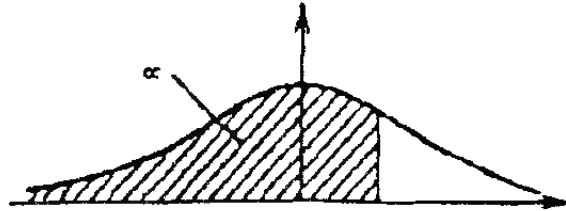
	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0.00	Z
										0.0013	-3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0017	0.0018	0.0019		-2.9
.0019	.0020	.0021	.0021	.0022	.0023	.0023	.0024	.0025	.0026		-2.8
.0026	.0027	.0028	.0029	.0030	.0031	.0032	.0033	.0034	.0035		-2.7
.0036	.0037	.0038	.0039	.0040	.0041	.0043	.0044	.0045	.0047		-2.6
.0048	.0049	.0051	.0052	.0054	.0055	.0057	.0059	.0060	.0062		-2.5
.0064	.0066	.0068	.0069	.0071	.0073	.0075	.0078	.0080	.0082		-2.4
.0084	.0087	.0089	.0091	.0094	.0096	.0099	.0102	.0104	.0107		-2.3
.0110	.0113	.0116	.0119	.0122	.0126	.0129	.0132	.0136	.0139		-2.2
.0143	.0146	.0150	.0154	.0158	.0162	.0166	.0170	.0174	.0179		-2.1
.0183	.0188	.0192	.0197	.0202	.0207	.0212	.0217	.0222	.0227		-2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287		-1.9
.0294	.0301	.0307	.0314	.0322	.0329	.0336	.0344	.0351	.0359		-1.8
.0367	.0375	.0384	.0392	.0401	.0409	.0418	.0427	.0436	.0446		-1.7
.0455	.0465	.0475	.0485	.0495	.0505	.0516	.0526	.0537	.0548		-1.6
.0559	.0571	.0582	.0594	.0606	.0618	.0630	.0643	.0655	.0668		-1.5
.0681	.0694	.0708	.0721	.0735	.0749	.0764	.0778	.0793	.0808		-1.4
.0823	.0838	.0853	.0869	.0885	.0901	.0918	.0934	.0951	.0968		-1.3
.0985	.1003	.1020	.1038	.1056	.1075	.1093	.1112	.1131	.1151		-1.2
.1170	.1190	.1210	.1230	.1261	.1271	.1292	.1314	.1335	.1357		-1.1
.1379	.1401	.1423	.1446	.1469	.1492	.1515	.1539	.1562	.1587		-1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841		-0.9
.1867	.1894	.1921	.1949	.1977	.2006	.2033	.2061	.2090	.2119		-0.8
.2148	.2177	.2206	.2236	.2266	.2296	.2327	.2358	.2389	.2420		-0.7
.2451	.2483	.2514	.2546	.2578	.2611	.2643	.2676	.2709	.2743		-0.6
.2776	.2810	.2843	.2877	.2912	.2946	.2981	.3015	.3050	.3085		-0.5
.3121	.3156	.3192	.3228	.3264	.3300	.3336	.3372	.3409	.3446		-0.4
.3483	.3520	.3557	.3594	.3632	.3669	.3707	.3745	.3783	.3821		-0.3
.3859	.3897	.3936	.3974	.4013	.4052	.4090	.4129	.4168	.4207		-0.2
.4247	.4286	.4325	.4364	.4404	.4443	.4483	.4522	.4562	.4602		-0.1
.4641	.4681	.4721	.4761	.4801	.4840	.4880	.4920	.4960	.5000		0.0
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359	
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753	
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141	
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517	
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879	
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224	
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549	
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852	
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8079	.8106	.8133	
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389	
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621	
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830	
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015	
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177	
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319	
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441	
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545	
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633	
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706	
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767	
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817	
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857	
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890	
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9915	
2.4	.9916	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936	
2.5	.9939	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952	
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964	
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974	
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981	
2.9	.9981	.9982	.9983	.9983	.9984	.9984	.9985	.9985	.9986	.9986	
3.0	0.9987										

جدول (2-2) توزیع χ^2 - Chi Square



n	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$	$\chi^2_{.001}$	$\chi^2_{.0005}$	$\chi^2_{.0001}$	$\chi^2_{.00005}$	$\chi^2_{.00001}$	$\chi^2_{.000005}$	$\chi^2_{.000001}$	$\chi^2_{.0000005}$	$\chi^2_{.0000001}$	$\chi^2_{.00000005}$	$\chi^2_{.00000001}$
1	.000	.000	.001	.004	.016	.064	.149	.455	1.07	1.64	2.71	3.84	5.02	6.63	7.88	
2	.010	.020	.051	.103	.211	.446	.713	1.39	2.41	3.22	4.61	5.99	7.38	9.21	10.6	
3	.077	.115	.215	.352	.584	1.00	1.42	2.37	3.66	4.64	6.25	7.81	9.35	11.3	12.8	
4	.207	.297	.484	.711	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	11.1	13.3	14.9	
5	.472	.554	.831	1.15	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.1	12.8	15.1	16.7	
6	.676	.872	1.24	1.64	2.20	3.07	3.83	5.35	7.23	8.56	10.6	12.6	14.4	16.8	18.5	
7	.989	1.24	1.69	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.0	14.1	16.0	18.5	20.3	
8	1.34	1.65	2.18	2.73	3.49	4.59	5.33	7.34	9.52	11.0	13.4	15.5	17.5	20.1	22.0	
9	1.73	2.09	2.70	3.33	4.17	5.38	6.39	8.34	10.7	12.2	14.7	16.9	19.0	21.7	23.6	
10	2.16	2.56	3.25	3.94	4.87	6.18	7.27	9.34	11.8	13.4	16.0	18.3	20.5	23.2	25.2	
11	2.60	3.05	3.82	4.67	5.58	6.99	8.15	10.3	12.9	14.6	17.3	19.7	21.9	24.7	26.8	
12	3.07	3.57	4.40	5.23	6.30	7.81	9.03	11.3	14.0	15.8	18.5	21.0	23.3	26.2	28.3	
13	3.57	4.11	5.01	5.89	7.04	8.63	9.93	12.3	15.1	17.0	19.8	22.4	24.7	27.7	29.8	
14	4.07	4.66	5.63	6.57	7.79	9.47	10.8	13.3	16.2	18.2	21.1	23.7	26.1	29.1	31.3	
15	4.60	5.23	6.26	7.26	8.55	10.3	11.7	14.3	17.3	19.3	22.3	25.0	27.5	30.6	32.8	
16	5.14	5.81	6.91	7.96	9.31	11.2	12.6	15.3	18.4	20.5	23.5	26.3	28.8	32.0	34.3	
17	5.70	6.41	7.59	8.67	10.1	12.0	13.5	16.3	19.5	21.6	24.8	27.6	30.2	33.4	35.7	
18	6.26	7.01	8.23	9.39	10.9	12.9	14.4	17.3	20.6	22.8	26.0	28.9	31.5	34.8	37.2	
19	6.83	7.63	8.91	10.1	11.7	13.7	15.4	18.3	21.7	23.9	27.2	30.1	32.9	36.2	38.6	
20	7.43	8.26	9.59	10.9	12.6	14.6	16.3	19.3	22.8	25.0	28.4	31.4	34.2	37.6	40.1	
21	8.03	8.90	10.3	11.6	13.2	15.4	17.2	20.3	23.9	26.2	29.6	32.7	35.5	39.0	41.4	
22	8.64	9.54	11.0	12.3	14.0	16.3	18.1	21.3	24.9	27.3	30.8	33.9	36.8	40.3	42.8	
23	9.26	10.2	11.7	13.1	14.8	17.2	19.0	22.3	26.0	28.4	32.0	35.2	38.1	41.6	44.2	
24	9.89	10.9	12.4	13.8	15.7	18.1	19.9	23.3	27.1	29.6	33.2	36.4	39.4	43.0	45.6	
25	10.5	11.5	13.1	14.6	16.5	18.9	20.9	24.3	28.2	30.7	34.4	37.7	40.6	44.1	46.9	
26	11.2	12.2	13.8	15.4	17.3	19.8	21.8	25.3	29.2	31.8	35.6	38.9	41.9	45.4	48.1	
27	11.8	12.9	14.6	16.2	18.1	20.7	22.7	26.3	30.3	32.9	36.7	40.1	43.2	47.0	49.6	
28	12.5	13.6	15.3	16.9	18.9	21.6	23.6	27.3	31.4	34.0	37.9	41.3	44.5	48.3	51.0	
29	13.1	14.3	16.0	17.7	19.8	22.5	24.6	28.3	32.5	35.1	39.1	42.6	45.7	49.6	52.3	
30	13.8	15.0	16.8	18.5	20.6	23.4	25.5	29.3	33.5	36.2	40.3	43.8	47.0	50.9	53.7	
40	20.7	22.1	24.4	26.5	29.0	32.3	34.9	39.3	44.2	47.8	51.8	55.8	59.3	63.7	66.8	
50	28.0	29.7	32.3	34.8	37.7	41.4	44.3	49.3	54.7	58.2	63.2	67.5	71.4	76.2	79.5	
60	35.5	37.5	40.3	43.2	46.5	50.6	53.8	59.3	65.2	69.0	74.4	79.1	83.3	88.4	91.0	

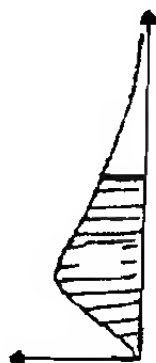
جدول (3-3) توزیع t



n	α = 0.6	0.75	0.9	0.95	0.975	0.99	0.995	0.9975	0.999	0.9995
1	0.323	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

جدول (4-4) توزيع F

	البسط/المقام	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	120	∞
90	1	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	60.7	61.2	61.7	62.3	62.8	63.1	63.3
95		161	200	216	225	230	234	237	239	241	242	244	246	248	250	252	253	254
97.5		648	800	864	900	922	937	948	957	963	969	977	983	993	1000	1010	1016	1020
99		4,050	5,000	5,400	5,670	5,860	5,930	5,950	5,980	6,070	6,080	6,110	6,160	6,210	6,260	6,340	6,370	6,400
99.5		16,200	20,000	21,600	22,500	23,100	23,400	23,700	23,900	24,100	24,200	24,400	24,600	24,800	25,000	25,200	25,400	25,500
90	2	8.51	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.46	9.47	9.48	9.49
95		18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5
97.5		38.5	39.0	39.2	39.3	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5
99		98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5
99.5		199	199	199	199	199	199	199	199	199	199	199	199	199	199	199	199	199
90	3	5.54	5.45	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.17	5.15	5.14	5.13
95		10.1	9.55	9.28	9.12	9.01	8.94	8.87	8.83	8.81	8.79	8.74	8.70	8.66	8.62	8.57	8.53	8.51
97.5		17.4	16.8	15.4	14.5	13.5	12.5	11.6	10.7	10.3	10.4	10.3	10.3	10.2	10.1	10.0	9.99	9.98
99		34.1	30.8	29.3	28.2	27.9	27.7	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.5	26.3	26.2	26.1
99.5		55.6	49.8	47.5	46.2	45.4	44.8	44.4	44.1	43.9	43.7	43.4	43.1	42.8	42.5	42.3	42.0	41.8
90	4	4.34	4.22	4.19	4.11	4.05	4.01	3.98	3.95	3.93	3.92	3.90	3.87	3.84	3.82	3.79	3.78	3.76
95		7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.69	5.64	5.61
97.5		12.2	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.46	8.36	8.31	8.26
99		21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.8	13.7	13.6	13.5
99.5		31.3	26.5	24.3	23.2	22.5	22.0	21.6	21.4	21.1	21.0	20.7	20.4	20.2	19.9	19.6	19.3	19.3
90	5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.17	3.14	3.12	3.11
95		6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.43	4.40	4.37
97.5		10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.63	6.52	6.43	6.33	6.23	6.12	6.07	6.02
99		16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.84	9.72	9.53	9.38	9.20	9.11	9.02
99.5		22.8	18.3	16.5	15.6	14.9	14.5	14.2	14.0	13.8	13.6	13.4	13.1	12.9	12.7	12.4	12.3	12.1
90	6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.80	2.76	2.74	2.72
95		5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.74	3.70	3.67
97.5		8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.07	4.96	4.90	4.83
99		13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.23	7.06	6.97	6.88
99.5		18.6	14.3	12.9	12.0	11.5	11.1	10.8	10.6	10.4	10.2	10.0	9.81	9.59	9.36	9.12	9.00	8.88
90	7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.73	2.72	2.70	2.67	2.63	2.59	2.56	2.51	2.49	2.47
95		5.39	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.38	3.30	3.27	3.23
97.5		8.01	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.36	4.23	4.20	4.14
99		12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.86	6.72	6.63	6.47	6.31	6.16	5.99	5.82	5.74	5.65
99.5		16.2	12.4	10.9	10.1	9.52	9.16	8.89	8.68	8.51	8.38	8.18	7.97	7.75	7.53	7.31	7.19	7.08
90	8	3.46	3.11	2.92	2.81	2.71	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.38	2.34	2.31	2.29
95		5.32	4.66	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.23	3.15	3.08	3.01	2.97	2.93
97.5		7.52	6.06	5.42	5.05	4.82	4.65	4.51	4.43	4.36	4.30	4.20	4.10	4.00	3.89	3.78	3.73	3.67
99		11.3	8.65	7.56	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.51	5.36	5.20	5.03	4.92	4.86
99.5		14.1	11.0	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01	6.81	6.61	6.40	6.18	6.04	5.95



90	1.14	2.81	2.69	2.61	2.31	2.47	2.44	1.47	2.38	2.34	2.30	2.23	2.21	2.19	2.16
91	3.12	2.86	3.63	3.48	3.37	3.23	3.18	3.14	3.07	3.01	2.94	2.84	2.79	2.73	2.71
95	7.21	6.99	4.72	4.68	4.32	4.10	4.03	3.96	3.87	3.77	3.67	3.54	3.45	3.39	3.33
99	10.6	6.99	6.42	6.04	5.80	5.47	5.25	5.24	5.11	4.94	4.81	4.63	4.49	4.40	4.31
995	13.4	8.72	7.96	7.47	7.13	6.69	6.54	6.42	6.23	6.03	5.83	5.62	5.41	5.30	5.19
90	3.29	2.73	2.61	2.52	2.44	2.38	2.33	2.22	2.28	2.24	2.20	2.18	2.11	2.09	2.06
95	4.96	4.10	3.48	3.33	3.22	3.07	3.02	2.98	2.91	2.84	2.77	2.70	2.62	2.59	2.54
975	6.94	4.83	4.47	4.24	4.07	3.93	3.78	3.72	3.62	3.52	3.42	3.31	3.20	3.14	3.08
99	10.0	7.56	6.53	6.04	5.39	5.06	4.94	4.85	4.71	4.56	4.41	4.23	4.04	4.00	3.91
995	12.8	8.08	7.34	6.87	6.34	5.62	5.97	5.85	5.66	5.47	5.27	5.07	4.86	4.75	4.64
90	2.18	2.61	2.48	2.39	2.33	2.24	2.21	2.19	2.13	2.10	2.06	2.01	1.96	1.93	1.90
91	4.73	3.49	3.28	3.11	3.00	2.85	2.80	2.73	2.69	2.62	2.54	2.47	2.38	2.34	2.30
95	6.53	5.10	4.47	4.12	3.89	3.73	3.54	3.37	3.28	3.18	3.07	2.96	2.85	2.79	2.72
99	9.33	6.93	5.63	5.04	4.64	4.50	4.39	4.30	4.16	4.01	3.84	3.70	3.54	3.43	3.36
995	11.8	8.31	7.23	6.07	5.76	5.33	5.30	5.09	4.91	4.72	4.53	4.33	4.12	4.01	3.90
90	3.07	2.70	2.49	2.37	2.31	2.12	2.09	2.04	2.02	1.97	1.92	1.87	1.82	1.79	1.76
91	4.54	3.29	3.06	2.90	2.79	2.64	2.59	2.54	2.46	2.40	2.33	2.25	2.16	2.11	2.07
95	6.20	4.77	4.13	3.88	3.61	3.29	3.13	3.04	2.96	2.86	2.76	2.64	2.52	2.46	2.40
99	8.68	6.36	5.42	4.83	4.32	4.00	3.87	3.70	3.63	3.52	3.37	3.21	3.04	2.94	2.87
995	10.8	7.70	6.48	5.37	5.07	4.67	4.54	4.42	4.25	4.07	3.88	3.69	3.48	3.37	3.28
90	2.97	2.59	2.38	2.25	2.09	2.04	1.94	1.94	1.89	1.84	1.79	1.74	1.68	1.64	1.61
91	4.35	3.10	2.87	2.71	2.80	2.51	2.39	2.33	2.23	2.20	2.12	2.04	1.95	1.90	1.84
95	5.87	4.46	3.86	3.51	3.13	2.91	2.84	2.77	2.68	2.57	2.46	2.35	2.27	2.16	2.09
99	8.10	5.85	4.94	4.43	3.87	3.56	3.44	3.37	3.23	3.09	2.94	2.78	2.61	2.52	2.42
995	9.94	6.99	5.82	5.17	4.47	4.09	3.96	3.83	3.68	3.50	3.32	3.12	2.92	2.81	2.69
90	2.88	2.49	2.38	2.14	1.98	1.83	1.81	1.82	1.77	1.72	1.67	1.61	1.54	1.50	1.46
91	4.17	3.12	2.92	2.69	2.33	2.23	2.21	2.16	2.09	2.01	1.93	1.84	1.74	1.68	1.62
95	5.37	4.18	3.39	3.23	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.07	1.94	1.87	1.79
99	7.54	5.39	4.51	4.02	3.47	3.17	3.07	2.98	2.84	2.70	2.55	2.39	2.21	2.11	2.01
995	9.18	6.35	5.24	4.62	3.95	3.58	3.45	3.34	3.18	3.01	2.82	2.63	2.42	2.30	2.18
90	2.79	2.39	2.18	2.04	1.87	1.77	1.74	1.71	1.66	1.60	1.54	1.48	1.40	1.37	1.39
91	4.00	3.15	2.76	2.51	2.17	2.10	2.04	1.99	1.92	1.84	1.73	1.63	1.53	1.47	1.39
95	5.29	3.93	3.34	3.01	2.53	2.41	2.33	2.27	2.17	2.06	1.94	1.82	1.67	1.58	1.48
99	7.08	4.98	4.13	3.64	3.12	2.82	2.72	2.63	2.50	2.35	2.20	2.03	1.84	1.73	1.60
995	8.49	5.80	4.73	4.14	3.49	3.13	3.01	2.90	2.74	2.57	2.39	2.19	1.96	1.83	1.69
90	2.73	2.35	2.13	1.99	1.82	1.72	1.68	1.65	1.60	1.54	1.48	1.41	1.32	1.26	1.19
91	3.92	3.07	2.68	2.45	2.09	2.02	1.96	1.91	1.83	1.73	1.66	1.53	1.43	1.35	1.25
95	5.15	3.80	3.21	2.89	2.38	2.30	2.22	2.16	2.05	1.94	1.81	1.69	1.51	1.43	1.31
99	6.63	4.79	3.93	3.46	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.84	1.66	1.53	1.38
995	8.18	5.54	4.50	3.92	3.28	2.93	2.81	2.71	2.54	2.37	2.19	1.98	1.73	1.61	1.43
90	2.71	2.30	2.08	1.94	1.77	1.67	1.63	1.60	1.55	1.49	1.42	1.34	1.24	1.17	1.00
91	3.84	3.00	2.60	2.37	2.10	1.94	1.88	1.83	1.73	1.61	1.47	1.40	1.32	1.22	1.00
95	5.07	3.69	3.12	2.79	2.29	2.19	2.11	2.05	1.94	1.81	1.71	1.46	1.36	1.27	1.00
99	6.93	4.61	3.72	3.02	2.64	2.51	2.42	2.32	2.18	2.04	1.88	1.70	1.47	1.32	1.00
995	7.86	5.30	4.28	3.32	3.09	2.74	2.62	2.52	2.36	2.19	2.00	1.79	1.53	1.36	1.00

المراجع

1. طرق التحليل الإحصائي . د. عبد العزيز فهمي هيكل - دار النهضة العربية للطباعة والنشر - بيروت - لبنان - 1998 .
2. الإحصاء التطبيقي في مجال البحوث التربوية والنفسية والاجتماعية والرياضية . د. مصطفى حسين باهي - مركز الكتاب الناشر - القاهرة - مصر - 1999 .
3. أساسيات الإحصاء وتطبيقاته . د. حسن محمد حسن محمد - دار المعرفة الجامعية - الإسكندرية - مصر - 1992 .
4. الإحصاء . د. أحمد عباده سرحان ، د. فاروق عبد العظيم أحمد - الطبعة الثانية - الإسكندرية - مصر - 1987 .
5. نظرية الاحتمالات . ب . غندينكو - ترجمة د. جمال الدباغ - دار مير - موسكو - 1990 .
6. مبادئ الإحصاء والاحتمالات . أنور اللحام ، محمد شفيق ياسين - دار الطباعة الحديثة - دمشق 1982 .
7. أساسيات علم الإحصاء . أوليف جين دوون - ترجمة د. عبد الرزاق الهوني وآخرون - منشورات جامعة الفاتح - طرابلس - ليبيا 1989 .

8. مبادئ في نظرية الاحتمالات . د. دسوقي مصطفى - دار النهضة العربية
- لبنان - بيروت 1986 .

9. مبادئ الإحصاء الاستنتاجي . د. فاروق البشتي وآخرون - الزاوية -
منشورات جامعة 1994 .

10. الطرق الإحصائية التطبيقية للمعينة . د. عبد الحميد عبد المجيد
البلداوي - منشورات جامعة السابع من ابريل . الطبعة الأولى 1995 .

11. مبادئ الإحصاء والاحتمالات . د. عبد العزيز فهمي هيكل - كلية
التجارة - جامعة الإسكندرية - جمهورية مصر العربية .

12. الإحصاء وتطبيقاته الهندسية . د.نعمة أحمد عمارة - وزارة التعليم
العالي والبحث العلمي - الموصل - العراق 1988.

13. مقدمة في الإحصاء الوصفي . د.مصطفى عبد المنعم الخواجة - كلية
التجارة - جامعة الإسكندرية - جمهورية مصر العربية .

14. مبادئ الإحصاء 2. ن.جاة رشيد الكيخيا - الشركة العربية للتنمية
والتجارة الدولية - طرابلس - ليبيا .

15. ملخصات شوم - نظريات ومسائل في الاحتمالات . د. سيمور ليبشتز .
ترجمة د.سامح داود - دار الرائد العربي - بيروت - لبنان 1984 .

16. Introduction to Probability and Statistics for Engineers
and Scientists S.M. Ross . J. Wiley Son 87 .